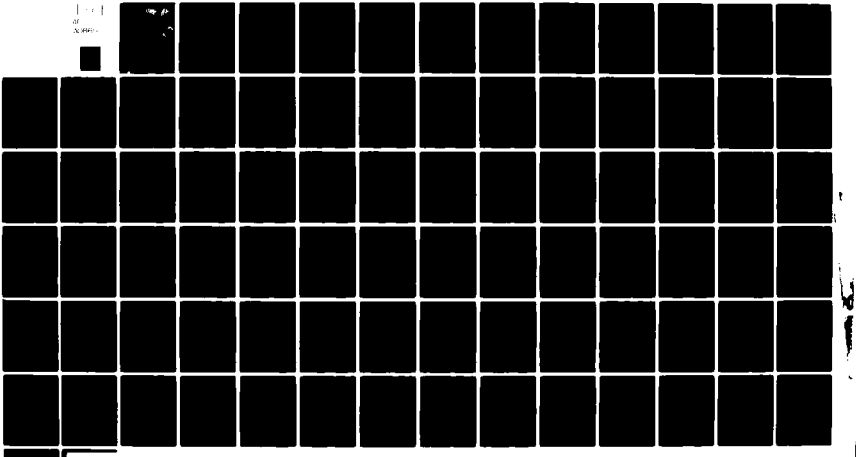
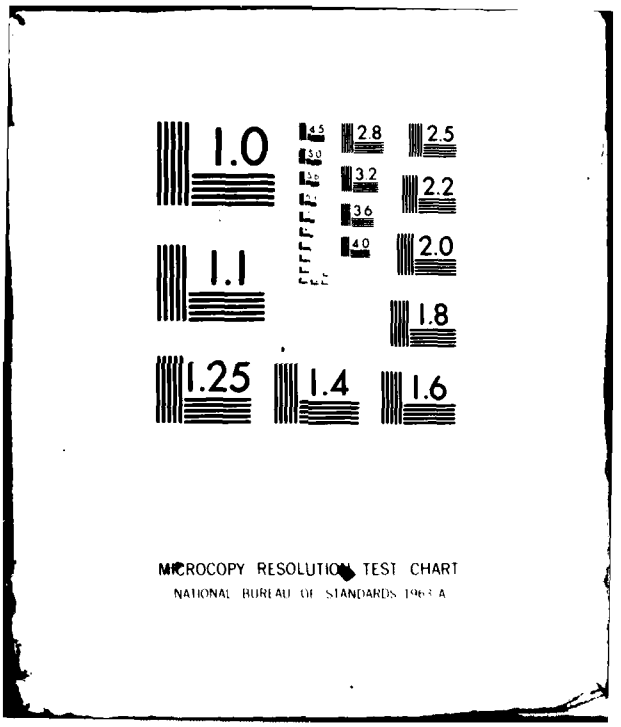
AD-A086 010 ROCHESTER UNIV N Y DEPT OF MECHANICAL AND AEROSPACE--ETC F/G 3/2
DYNAMICAL PHENOMENA IN SUNSPOTS.(U)
FEB 80 J H THOMAS, A CLARK, M A SCHEUER F19628-77-C-0079
UNCLASSIFIED AFGL-TR-80-0064 NL



END
DATE FILMED
8-80
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(19) REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (18) AFGL-TR-80-0064	2. GOVT ACCESSION NO. AD-A086 010	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) DYNAMICAL PHENOMENA IN SUNSPOTS	5. TYPE OF REPORT & PERIOD COVERED Final Report (period 1 October 1976 - 30 September 1979)
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) John H. Thomas, Alfred Clark, Jr., Mark A. Scheuer	8. CONTRACT OR GRANT NUMBER(s) (15) F19628-77-C-0079
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Mechanical & Aerospace Sciences University of Rochester Rochester, New York 14627	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (16) 61102F 2311G3AC (17) 63
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Geophysics Laboratory Hanscom AFB, Massachusetts 01731 Monitor/ Donald Neidig /PHS	12. REPORT DATE 15 February 1980
	13. NUMBER OF PAGES 82
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (11) 15 Feb 80 (12) 85	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(9) Final rept. 1 Oct 76-30 Sep 79.

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Sunspots
Umbral Oscillations
Magneto-atmospheric waves
Solar cycle

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The reflection of upward and downward propagating Alfvén waves in sunspots is studied, in order to assess the possibility of cooling by Alfvén waves. Wave reflection is studied by means of a three-layer model of the umbral atmosphere. The results show very strong downward reflection of Alfvén waves in the photosphere and low chromosphere, but only weak upward reflection in the convection zone. Further study of more realistic magneto-atmospheric waves (including the effects of compression and buoyancy) (over

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 40003.8 next page

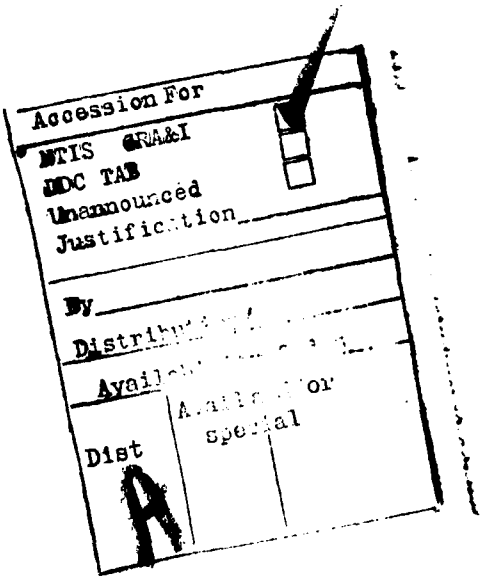
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

cont. shows strong upward reflection in the convection zone as well as strong downward reflection in the photosphere and low chromosphere. These results tend to rule out significant sunspot cooling by waves.

A study of a simple thermal model of a sunspot, based on the concept of partial inhibition of convection, shows that the inhibition mechanism can yield acceptable distributions of surface temperature. The results of this model also show that: (i) the edge of the umbra is sharp, even for deep spots; (ii) deep spots produce weak bright rings, but shallow spots produce intense bright rings in conflict with observations; and (iii) only a shallow surface layer of the sunspot is cool, the rest being warmer than the surroundings.

Umbral oscillations in sunspots are studied and identified as a resonant response of the umbral atmosphere to forcing by oscillatory convection in the subphotosphere. The full linearized equations for magneto-atmospheric waves are solved numerically for a detailed model of the umbral atmosphere. Resonant "fast" modes are found, the lowest mode having a period of 153 s, typical of umbral oscillations. A small amount of wave energy leaks into the corona in the form of an acoustic wave along the ~~magnetic field lines~~.

It is suggested that the Sun's radius and surface temperature vary with the solar cycle due to the variation of total magnetic buoyancy in the convection zone over the cycle of the solar dynamo. This mechanism might produce the drop in surface temperature with increasing solar activity observed by Livingston (1978), with little or no change in luminosity.



PERSONNEL

Air Force Contract F 19628-77-C-0079

John H. Thomas, Principal Investigator

Alfred Clark, Jr., Faculty Associate

Mark A. Scheuer, Research Assistant

CONTENTS

		Page
1.	Summary	5
2.	Thomas, J.H. 1978, The reflection of Alfvén waves and the cooling of sunspots, Astrophys. J. 225, 275-280. (Reprint)	11
3.	Clark, A., Jr. 1979, Thermal models of sunspots, Solar Phys. 62, 305-330. (Reprint)	17
4.	Thomas, J.H. 1979, Variations of the Sun's radius and temperature due to magnetic buoyancy, Nature 280, 662-663. (Reprint)	43
5.	Scheuer, M.A. and Thomas, J.H. 1980, Umbral oscillations as resonant modes of magneto-atmospheric waves. (Preprint)	45

SUMMARY

The research under this contract has focused upon dynamical and thermal phenomena in sunspots. The research may be roughly divided into three categories: (i) a study of the possible mechanisms for cooling a sunspot, including wave cooling and inhibition of convection; (ii) a study of resonant modes of magneto-atmospheric waves in a sunspot umbra, as an explanation for umbral oscillations; and (iii) a suggestion that the Sun's radius and surface temperature may vary over the solar cycle due to changes in total magnetic buoyancy in the convection zone. The last topic, while somewhat outside of the scope of the initial proposal, nevertheless arose in the course of this research.

The question of the mechanism that causes a sunspot to be cooler than its surroundings has proved to be quite controversial. The traditional explanation has been based on the inhibition of convective heat transport by the strong magnetic field in the sunspot (Biermann 1941). An alternative mechanism, based on cooling by an outward flux of Alfvén waves (or other magnetohydrodynamic waves), has been advocated, most recently by Parker (1974, 1975). Both of these mechanisms have been investigated further under this contract, with the conclusion that the inhibition mechanism is by far the more likely explanation.

Thomas (1978; section 2 of this report) investigated the reflection of Alfvén waves in the umbral atmosphere. He considered the reflection of waves propagating upward and downward in the umbra from a generating source (overstable convection) in a shallow layer just below the umbral photosphere. The study used a three-layer model of the umbral atmosphere that reproduces

all of the essential features of the vertical structure in the convection zone, photosphere, chromosphere, transition region, and corona. The results show very strong downward reflection of Alfvén waves propagating upward into the photosphere and low chromosphere. This result is in good agreement with the observations of Beckers (1976) and Beckers and Schneeberger (1977), which show a large drop in wave energy density from the photosphere to the high chromosphere. The observations put an upper limit on the upward wave energy flux that is several orders of magnitude below the "missing" flux of a sunspot. Thus, both theory and observation here tend to rule out substantial cooling by upward-propagating Alfvén waves.

On the other hand, Thomas (1978) found only very weak upward reflection of Alfvén waves propagating downward into the convection zone. This left open the possibility of cooling by downward propagating waves. Several difficulties associated with cooling by downward-propagating waves were noted, however. In particular, it was speculated that if one were to study more realistic wave motions, including the effects of compression and buoyancy (rather than the pure Alfvén mode), then one would find much stronger upward reflection in the convection zone due to the rising temperature (and sound speed) with depth. Further work by Scheuer and Thomas (1980; section 5 of this report) confirmed this expectation. Using the same model umbral atmosphere, but considering the full magneto-atmospheric wave equations (with nonzero horizontal wavenumber), Scheuer and Thomas found strong upward reflection in the convection zone, due primarily to the increasing sound speed with depth. This strong reflection holds true even for waves with horizontal wavelengths comparable to the spot diameter. Thus, cooling by downward-propagating waves seems to be ruled out too. Our conclusion is

that the wave cooling mechanism may be ruled out on theoretical and observational grounds.

In order to assess the possibility of sunspot cooling by the mechanism of inhibition of convection, Clark (1979, Section 3 of this report) studied simple thermal models of sunspots which involve solutions to the steady heat equation with an appropriate distribution of thermal conductivity. Similar studies had been made by Parker (1974), Eschrich and Krause (1977), and Spruit (1977). Clark chose a simple model for the sunspot which allowed him to investigate the dependence of the solution on the various spot parameters, such as the depth of region of inhibition. He found that his model sunspot, like those of Eschrich and Krause and of Spruit, produced surface temperature distributions very much like that observed in sunspots. Detailed results from Clark's model include the following: (i) the edge of the umbra is sharp, even for deep spots (depth of inhibition region $\gtrsim$ diameter); (ii) deep spots produce weak bright rings, whereas shallow spots produce intense bright rings (which aren't observed); and (iii) only a thin surface layer, about one temperature scale height in thickness, is cool, with the deeper parts of the spot somewhat hotter than the surroundings. The three thermal models of sunspots due to Eschrich and Krause (1977), Spruit (1977), and Clark (1979) each use a different specification for the thermal conductivity and for the spot geometry, and yet all three produce acceptable surface temperature distributions. This lends considerable support for the inhibition mechanism.

In further work on thermal models of sunspots, Clark has studied the effects of a depth-dependent thermal conductivity, first discussed by Spruit (1977). Using some simple analytical solutions, Clark has shown that the horizontal spreading of the heat flux depends on both the variable conductivity

and the boundary condition at the bottom of the convection zone. As a particular example, consider a steady heat source at depth z_0 below the top of the convection zone, and take the convection zone to be infinitely deep. At the top of the convection zone assume a radiative boundary condition. If the thermal conductivity is constant, all of the heat emitted by the source flows out the top, as one would expect. If, however, the thermal conductivity K varies with depth z, the result can be very different. For the conductivity law suggested by Spruit, namely $K \propto (1+\alpha z)^2$ with $\alpha^{-1} = 127$ km, much of the heat emitted by the source flows off to infinity rather than leaving through the top. In fact, the fraction of the heat flux leaving through the top is

$$\frac{1}{(1+\alpha z_0)(1+\alpha h)} ,$$

where h is the superadiabatic temperature scale height at the surface. Thus for deep disturbances $(\alpha z_0 >> 1)$ very little of the released heat ever reaches the surface. This result shows that the boundary condition at the bottom of the convection zone is likely to be very important.

Scheuer and Thomas (1980; Section 5 of this report) have presented a detailed theory of umbral oscillations, identifying them as the lowest resonant mode of fast magneto-atmospheric wave in the umbral atmosphere, excited by overstable convection in the subphotosphere. Their theory is similar to that of Uchida and Sakurai (1975), except that Scheuer and Thomas use the full magneto-atmospheric wave equations rather than the "quasi-Alfvén" approximation of Uchida and Sakurai. Numerical solutions of the full wave equations are obtained for forced and free oscillations in a three-layer model of a sunspot umbra. Several new features of the wave motion arise from this study, leading

to a new understandg of umbral oscillations. The resonant mode is trapped primarily by the increasing Alfvén speed upward into the chromosphere and by the increasing sound speed downward into the convection zone. The downward reflection is not complete, however; a small fraction of the total energy escapes to large heights by converting into the form of a pure acoustic wave along the vertical magnetic field lines. The change in character of the wave with height, from Alfvén-like to acoustic-like, is one of the more interesting features predicted by the full equations. It is also found that the increasing sound speed in the convection zone, not the increasing density (as proposed by Uchida and Sakurai), provides the necessary upward reflection for the trapped mode. For numerical values of the model parameters based on a typical sunspot, the lowest resonant fast mode has a period of 153 s, typical of umbral oscillations.

On another topic, Thomas (1979, Section 4 of this report) has suggested that the solar radius and surface temperature may vary with the solar cycle, due to changes in the total magnetic buoyancy in the convection zone. This idea was prompted by the observations of Livingston (1978), which show a decrease in surface temperature of the Sun with increasing solar activity. Livingston interprets this to imply a corresponding decrease in luminosity, but Thomas points out that it could also be due to a small increase in the solar radius with little or no change in luminosity. The mechanism proposed for the expansion with increasing solar activity is based on increasing magnetic buoyancy in the convection zone as the solar dynamo moves toward solar maximum. Rough estimates of the difference in magnetic flux in the convection zone between solar maximum and minimum indicate that this mechanism may be important. The historical record of observations of the solar radius

is quite inconsistent, but there is evidence of variations of the solar radius with the solar cycle, with maximum radius at maximum activity. Accurate monitoring of the solar radius over a solar cycle is needed to evaluate this idea and to provide needed information for the relation between solar luminosity, temperature, and size.

References

Beckers, J.M. 1976, Astrophys. J. 203, 739.

Beckers, J.M., and Schneeberger, T.J. 1977, Astrophys. J. 215, 356.

Biermann, L. 1941, Vierteljahrsschr. Astron. Gesellsch. 76, 194.

Clark, A., Jr. 1979, Solar Phys. 62, 305.

Eschrich, K.-O., and Krause, F. 1977, Astron. Nach. 298, 1.

Livingston, W.C. 1978, Nature 272, 340.

Parker, E.N. 1974, Solar Phys. 36, 249; 37, 127.

Parker, E.N. 1975, Solar Phys. 40, 275; 40, 291.

Scheuer, M.A., and Thomas, J.H. 1980, submitted to Solar Phys.

Spruit, H.C. 1977, Solar Phys. 55, 3.

Thomas, J.H. 1978, Astrophys. J. 225, 275.

Thomas, J.H. 1979, Nature 280, 662.

Uchida, Y., and Sakurai, T. 1975, Publ. Astron. Soc. Japan 27, 259.

THE ASTROPHYSICAL JOURNAL, 225:275–280, 1978 October 1

THE REFLECTION OF ALFVÉN WAVES AND THE COOLING OF SUNSPOTS

JOHN H. THOMAS
Department of Mechanical and Aerospace Sciences and C. E. Kenneth Mees Observatory, University of Rochester
Received 1978 March 7; accepted 1978 April 6

ABSTRACT

As one means of evaluating the possibility that sunspots are cooled by a flux of Alfvén waves, the reflection of vertically propagating Alfvén waves in a three-layer model of a sunspot umbra is studied. The results show strong downward reflection of Alfvén waves in the photosphere and low chromosphere, with very little wave energy penetrating as high as the corona. This is in agreement with recent observations. The model umbra also shows very weak upward reflection of Alfvén waves propagating downward into the convection zone. The results suggest that, if sunspots are indeed cooled by Alfvén waves, these waves must escape downward into the solar interior.

Subject headings: hydromagnetics — Sun: sunspots

I. INTRODUCTION

The mechanism that causes a sunspot to be cooler than its surroundings is still not understood. Biermann's (1941) suggestion that the cause is the inhibition of convection by the intense magnetic field in the sunspot has often been accepted but has never developed into a full-fledged theory. Parker (1974*a*, *b*, 1975*a*, *b*) recently revived an alternative suggestion, based on earlier work by others (Danielson 1965; Musman 1967; Savage 1969; Moore 1973), that the cooling is due to a flux of Alfvén waves (or other magneto-atmospheric waves) generated by overstable convection in the umbral subphotosphere (see also Roberts 1976). It is possible, of course, that both mechanisms operate simultaneously; indeed, it is the inhibiting effect of the magnetic field on convective motions that leads to the overstability that supposedly generates the Alfvén waves.

Parker's ideas have led to a lively controversy over several points. Cowling (1976) has objected to the high thermal efficiency required for the "refrigeration" of a sunspot by a flux of waves (see the rebuttal by Parker 1977). By means of a simple thermal model, Parker (1974*a*) argued that the inhibition mechanism would not produce the observed surface temperature distribution of a sunspot. But other thermal models (Eschrich and Krause 1977; Spruit 1977) have produced acceptable sunspots, and Clark (1978) has shown that this point hinges on what one assumes the depth of a sunspot to be (see also Wilson 1971). Boruta (1977) has argued that the formation of a sunspot is more readily explained by the wave-cooling mechanism. Another question concerns the intrinsic stability of a fully developed sunspot (Parker 1975*b*; Meyer, Schmidt, and Weiss 1977), which may depend on the cooling mechanism.

In the present paper we deal with a specific problem associated with the possible cooling by Alfvén waves: the escape of Alfvén waves may be limited by reflections in the strong vertical density gradients in the sunspot. The observations of Beckers (1976) and Beckers and Schneeberger (1977), taken together, indicate a strong downward reflection of Alfvén waves in the sunspot atmosphere. Beckers (1976) puts an upper bound of 2.5×10^{10} ergs cm^{-2} s^{-1} on the upward energy flux of Alfvén waves in the umbral photosphere, and Beckers and Schneeberger (1977) put an upper bound of 4×10^{7} ergs cm^{-2} s^{-1} on the upward energy flux of Alfvén waves in the corona above a sunspot umbra. The latter flux limit is well below the "missing" flux of 5×10^{10} ergs cm^{-2} s^{-1} needed to cool a sunspot. These observations leave open the possibility that the Alfvén waves escape downward into the solar interior.

If we assume a purely vertical, uniform umbral magnetic field $\boldsymbol{B}_0 = (0, 0, B_0)$, in a Cartesian coordinate system (x, y, z) with z upward, and assume purely horizontal motions with velocity $\boldsymbol{u} = [u(z, t), 0, 0]$, then we obtain the equation

$$\frac{\partial^2 u}{\partial t^2} = v_A^2(z) \frac{\partial^2 u}{\partial z^2} \tag{1}$$

describing vertically propagating pure Alfvén waves in the umbra, with Alfvén speed

$$v_A(z) = [B_0^2/4\pi\mu\rho(z)]^{1/2}$$

that varies due to the varying density $\rho(z)$. If we assume an oscillatory solution of the form $u(z, t) = \hat{u}(z) \exp(i\omega t)$, equation (1) becomes

$$\frac{d^2\hat{u}}{dz^2} + \frac{\omega^2}{v_A^2(z)} \hat{u} = 0 . \tag{2}$$

For a specified density distribution $\rho(z)$, we can solve equation (2) to determine effective reflection coefficients for Alfvén waves in the atmosphere. This type of problem is important in the study of waves in layered media (see Brekhovskikh 1960), and is analogous, for example, to the problem of reflection

of light in a medium in which the index of refraction varies in the direction of propagation.

Geronicolas (1977) has considered this problem with the density distribution[1]

$$\rho(z) = \rho_0[1 - (\rho_1/\rho_0)\tanh kz] \qquad (3)$$

representing the sunspot, where ρ_0, ρ_1, and k are parameters. He chose values of the parameters to match $\rho(z)$ with Spruit's (1974) model of the convection zone. The level $z = 0$ corresponds to the base of the umbral photosphere. Geronicolas found only a slight net downward reflection of Alfvén waves in this atmosphere. (To be precise, he found that if Alfvén waves are emitted upward and downward with equal intensity at $z = 0$, then the net reflection is such that the upward energy flux is 46% of the total, instead of 50% as in the case of no net reflection.) If we take this result together with the observations discussed above, we would conclude that Alfvén waves are not cooling sunspots.

Objections can be made to Geronicolas's density distribution (3), however. It gives a good fit to densities in detailed umbral models only over a very limited range of height in and near the photosphere. The density (3) approaches uniform values a short distance above and below the photosphere, whereas the true solar density continues to vary strongly at all levels above and below the surface. The net effect of this inconsistency is that Geronicolas's model severely underestimates the downward reflection of Alfvén waves in a sunspot. Here we propose a more realistic, three-layer model of the vertical density distribution in a sunspot umbra, and show that it leads to strong downward reflection of Alfvén waves.

II. THE MODEL UMBRA

The calculation of reflection coefficients for waves in a variable atmosphere is a case where approximate methods, such as the WKB method, usually fail. For this reason it is desirable to have a model umbra that is simple enough to allow exact solution of the wave equation (2), but detailed enough to give a fairly accurate representation of the vertical sunspot structure.

We adopt as our model umbra an atmosphere with a uniform vertical magnetic field and the three-layer temperature distribution shown in Figure 1. Layer 2 is isothermal at temperature T_0 and represents the umbral photosphere and chromosphere. The density in layer 2 is thus given by

$$\rho(z) = \rho_0 \exp(-z/H_0), \quad 0 \le z \le z_t, \qquad (4)$$

where $H_0 = RT_0/g$ is the density scale height in this layer. Layer 3, which represents the corona, is isothermal at temperature T_1 ($> T_0$), and the transition

[1] The formula for $\rho(z)$ has been changed here to agree with our present sign convention; Geronicolas took the z-axis to point downward. The problem of Alfvén wave reflection with the same density distribution, (3), has been considered by Adam (1976) in a different context.

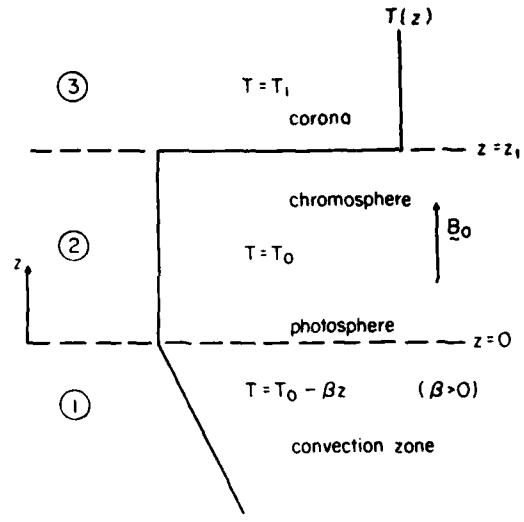


FIG. 1.—Schematic diagram of the three-layer temperature distribution in the model umbra.

region is represented as a discontinuity in temperature (and density) at $z = z_t$. Pressure must be continuous across $z = z_t$, however, and thus the density in layer 3 is given by

$$\rho(z) = \rho_0 \exp(-z_t/H_0)\left(\frac{T_0}{T_1}\right) \exp[-(z - z_t)/H_1], \quad z > z_t, \qquad (5)$$

where $H_1 = RT_1/g$ is the scale height in this layer.

Layer 1, which represents the umbral convection zone, is assumed to have a linear temperature distribution of the form $T(z) = T_0 - \beta z$, with $\beta > 0$. The value of β may be chosen equal to or slightly greater than the adiabatic gradient $\beta_a = g/c_p$. The corresponding density distribution in layer 1 is then

$$\rho(z) = \rho_0\left(1 - \frac{\alpha z}{H_0}\right)^{(1-\alpha)/\alpha}, \quad z < 0, \qquad (6)$$

where $\alpha = R\beta/g$ is a dimensionless measure of the temperature gradient β.

Layers 2 and 3 are the same as the two-layer model umbra studied by Uchida and Sakurai (1975) in connection with umbral oscillations. The same temperature structure as in layers 2 and 3, but with a vertical dipole magnetic field, was studied by Uchida and Kaburaki (1974) in connection with the heating of the chromosphere and corona above sunspots. We shall see here that layer 3 (the corona) is unimportant as far as the reflection of Alfvén waves that might be cooling a sunspot is concerned.

III. DOWNWARD REFLECTION OF ALFVÉN WAVES

We imagine that Alfvén waves are generated in a thin layer near $z = 0$ in the model umbra, with equal intensity upward and downward, and we then ask

how much of the wave energy is reflected downward or upward. First we consider the downward reflection from layers 2 and 3. To isolate this aspect of the problem, we imagine for the moment that the region $z < 0$ is homogeneous with density ρ_0 and that there is a train of plane Alfvén waves of fixed frequency ω and unit amplitude incident upon $z = 0$ from below. We then write, for $z < 0$,

$$u(z, t) = \exp\left[i\omega\left(t - \frac{z}{v_{A_0}}\right)\right] + U \exp\left[i\omega\left(t + \frac{z}{v_{A_0}}\right)\right], \tag{7}$$

where $v_{A_0} = (B_0^2/4\pi\mu\rho_0)^{1/2}$ is the Alfvén speed at $z = 0$. The first term on the right-hand side of equation (7) is an incident (upward propagating) wave of unit amplitude. The second term on the right-hand side is the reflected (downward propagating) wave of undetermined amplitude U. The reflected amplitude is determined by matching the solution (7) to the appropriate solution in layers 2 and 3. The proper matching conditions at the interfaces ($z = 0$ and $z = z_t$) are that the velocity $\hat{u}$ and the x-component of the magnetic field perturbation be continuous across the interface; the latter condition is equivalent to requiring $d\hat{u}/dz$ to be continuous across the interface. Once U is determined, the downward reflection coefficient for energy, R_d, is given by $R_d = |U|^2$.

First, let us ignore layer 3 and assume that the density distribution of layer 2 (eq. [4]) extends to $z = +\infty$. (The reason for doing this will become clear in a moment.) As first shown by Ferraro (1954; see also Ferraro and Plumpton 1958), the wave equation (2) with density distribution (4) can be transformed into Bessel's equation,

$$\zeta^2 \frac{d^2\hat{u}}{d\zeta^2} + \zeta \frac{d\hat{u}}{d\zeta} + \zeta^2\hat{u} = 0, \tag{8}$$

with the change of independent variable $\zeta = a \exp(-z/2H_0)$, where $a = 2H_0\omega/v_{A_0}$. Note that the parameter a can also be written as $a = 4\pi H_0/\lambda$, where $\lambda = 2\pi v_{A_0}/\omega$ is the wavelength of the incident Alfvén wave; thus a is a nondimensional wavenumber. The general solution of equation (8) is written as usual as

$$\hat{u}(\zeta) = AJ_0(\zeta) + BY_0(\zeta). \tag{9}$$

We require the solution to be finite as $z \to +\infty$ (i.e., outgoing wave as $z \to +\infty$), so $B = 0$. Then, matching $\hat{u}$ and $d\hat{u}/dz$ in solutions (7) and (9) at $z = 0$ ($\zeta = a$), we obtain

$$A = \frac{2}{J_0(a) + iJ_1(a)}, \tag{10}$$

and

$$U = \left[\frac{J_1(a) + iJ_0(a)}{J_1(a) - iJ_0(a)}\right]. \tag{11}$$

Thus, $R_d = |U|^2 = 1$ for any value of a, and there is total downward reflection of Alfvén waves of any wavelength. This result should not be surprising; any density distribution with $\rho \to 0$ as $z \to +\infty$ will give $R_d = 1$. Of more importance here is the fact that, for wavelengths expected for Alfvén waves associated with sunspot cooling, the reflection occurs quite low in the umbral atmosphere. The Alfvén waves do not propagate far enough upward to cool the spot.

A reasonable estimate of the scale height in layer 2 is $H_0 = 100$ km corresponding roughly to $T_0 = 4500$ K. Beckers (1976) estimates that upward-propagating Alfvén waves that might be cooling a sunspot would have wavelengths in the range 15–40 km. For these wavelengths, $a \geq 10\pi$, and we can use the asymptotic forms

$$J_0(a) = \left(\frac{2}{\pi a}\right)^{1/2} \cos\left(a - \frac{\pi}{4}\right),$$

$$J_1(a) = \left(\frac{2}{\pi a}\right)^{1/2} \sin\left(a - \frac{\pi}{4}\right).$$

Then equations (9) and (10) give

$$|\hat{u}| = (2\pi a)^{1/2} J_0(\zeta).$$

As $z \to +\infty (\zeta \to 0)$, $J_0(\zeta) \to 1$, and the wave amplitude approaches a uniform value $(2\pi a)^{1/2}$, while the kinetic energy density $\rho|\hat{u}|^2$ decreases rapidly because of the exponentially decreasing mass density. (The magnetic energy density also decreases exponentially with increasing z [Ferraro 1954].) We can write

$$\rho|\hat{u}|^2 = 2\pi a \rho J_0^2(\zeta) < 2\pi a \rho_0 \exp(-z/H_0),$$

and thus most of the wave energy is reflected downward within the first few scale heights above $z = 0$. A reasonable estimate of the height of the transition region is $z_t = 2000$ km. Since this is 20 scale heights above $z = 0$ in our model, only a very small fraction of the wave energy penetrates up to the transition region, and the effect of the corona (layer 3) is unimportant as far as the reflection of the Alfvén waves is concerned. For completeness the solution with layer 3 included is presented in the Appendix. The downward reflection coefficient is again unity, and the results of the present section are shown to be valid in the appropriate limit.

Note that the solution $\hat{u}(\zeta) \propto J_0(\zeta)$ oscillates with increasing z for $z > 0$ ($\zeta < a$) before becoming essentially uniform as $z \to +\infty$ ($\zeta \to 0$). Since, for example, $J_0(0.4) = 0.96$, the wave amplitude reaches 96% of its asymptotic value at height $z_0 = -2H_0 \log(0.4/a)$. The height z_0 is greater for waves of shorter wavelength (larger a); waves of shorter wavelength (higher frequency) propagate farther in the direction of increasing Alfvén speed. But even for a wavelength as short as 10 km ($a = 40\pi$), the height $z_0 = 1150$ km is well below the transition zone. For $z > z_0$, the wave amplitude is essentially constant, and the energy density continues to decrease exponentially in proportion to the mass density. The result that the strong downward reflection of Alfvén waves occurs very low in the sunspot atmosphere helps to justify our initial

assumption of a purely vertical magnetic field in the umbra.

In the observations of Beckers (1976) and Beckers and Schneeberger (1977), the directly observed quantity is the rms velocity of Alfvén waves causing the nonthermal line broadening. The values they obtain for the rms velocity are roughly 1.5 km s^{-1} for the photosphere and 7.5 km s^{-1} for the corona. This small observed change in rms velocity with height is consistent with the solution presented here.

IV. UPWARD REFLECTION OF ALFVÉN WAVES

To isolate the problem of determining the upward reflection coefficient in layer 1, we imagine the region $z > 0$ to be homogeneous with density ρ_0 and consider a train of plane Alfvén waves of fixed frequency incident upon $z = 0$ from above. We write, for $z > 0$,

$$u(z, t) = \exp\left[i\omega\left(t + \frac{z}{v_{A_0}}\right)\right] + V\exp\left[i\omega\left(t - \frac{z}{v_{A_0}}\right)\right]. \tag{12}$$

The first term on the right-hand side of equation (12) is an incident (downward propagating) wave of unit amplitude, and the second term on the right-hand side is the reflected (upward propagating) wave of undetermined amplitude V.

The basic wave equation (2) can be solved exactly in layer 1, where the density is given by equation (6). If we let

$$\eta = \frac{a}{\alpha + 1}\left(1 - \frac{\alpha z}{H_0}\right)^{(\alpha+1)/2\alpha},$$

where $a = 2H_0\omega/v_{A_0}$ as before, then the general solution of equation (2) in layer 1 can be expressed as (Watson 1966)

$$\hat{u}(\eta) = \eta^{\nu}[AH_{\nu}^{(1)}(\eta) + BH_{\nu}^{(2)}(\eta)], \tag{13}$$

where $\nu = \alpha/(\alpha + 1)$ and where $H_{\nu}^{(1)}$ and $H_{\nu}^{(2)}$ are Hankel functions of the first and second kinds of order ν. For $z \to -\infty$ ($\eta \to +\infty$), the principal asymptotic forms of the Hankel functions are

$$H_{\nu}^{(1)}(\eta) \sim \left(\frac{2}{\pi\eta}\right)^{1/2}\exp\left\{i\left[\eta - \left(\frac{\nu}{2} + \frac{1}{4}\right)\pi\right]\right\},$$

$$H_{\nu}^{(2)}(\eta) \sim \left(\frac{2}{\pi\eta}\right)^{1/2}\exp\left\{-i\left[\eta - \left(\frac{\nu}{2} + \frac{1}{4}\right)\pi\right]\right\}.$$

Thus, as $z \to -\infty$ ($\eta \to +\infty$), we see that $H_{\nu}^{(1)}$ represents a wave in the $+z$ ($-\eta$) direction and $H_{\nu}^{(2)}$ represents a wave in the $-z$ ($+\eta$) direction. Here we want only an outgoing wave as $z \to -\infty$, so we must take $A = 0$ in equation (13).

If we now match $\hat{u}$ and $d\hat{u}/dz$ in solutions (12) and (13) at $z = 0$, we obtain

$$B = 2\Big/\left\{\left(\frac{a}{\alpha+1}\right)^{\nu}\left[H_{\nu}^{(2)}\left(\frac{a}{\alpha+1}\right) + iH_{\nu-1}^{(2)}\left(\frac{a}{\alpha+1}\right)\right]\right\}, \tag{14}$$

and

$$V = \left[iH_{\nu}^{(2)}\left(\frac{a}{\alpha+1}\right) + H_{\nu-1}^{(2)}\left(\frac{a}{\alpha+1}\right)\right]\Big/\left[iH_{\nu}^{(2)}\left(\frac{a}{\alpha+1}\right) - H_{\nu-1}^{(2)}\left(\frac{a}{\alpha+1}\right)\right]. \tag{15}$$

The Hankel functions are complex valued, so $|V| \neq 1$ here. To evaluate V we must first choose an appropriate value of α, the dimensionless temperature gradient in layer 1. To avoid tedious computation we shall choose values of α such that $\nu = \alpha/(\alpha + 1)$ and $\nu - 1 = -1/(\alpha + 1)$ are among the fractional orders for which tables of Bessel functions are available. The values $\alpha = \frac{1}{2}$ (giving $\nu = \frac{1}{3}$, $\nu - 1 = -\frac{2}{3}$) and $\alpha = \frac{1}{3}$ (giving $\nu = \frac{1}{4}$ and $\nu - 1 = -\frac{3}{4}$) seem most appropriate. We can write

$$\alpha = \frac{R\beta}{g} = \left(\frac{c_p - c_v}{c_p}\right)\frac{c_p\beta}{g} = \left(\frac{\gamma - 1}{\gamma}\right)\frac{\beta}{\beta_s},$$

where $\beta_s = g/c_p$ is the adiabatic temperature gradient. Thus,

$$\frac{\beta}{\beta_s} = \left(\frac{\gamma}{\gamma - 1}\right)\alpha. \tag{16}$$

We can place bounds on α as follows. We want $\beta/\beta_s \geq 1$, which requires $\alpha \geq (\gamma - 1)/\gamma$. Most models of the umbral subphotosphere show density increasing with depth more rapidly than a linear variation, and we see from equation (6) that this requires $\alpha \leq \frac{1}{2}$. Thus, we have $(\gamma - 1)/\gamma \leq \alpha \leq \frac{1}{2}$. The lower bound equals $\frac{2}{5}$ for $\gamma = \frac{5}{3}$ or $\frac{1}{2}$ for $\gamma = \frac{3}{2}$. Thus, the two values $\alpha = \frac{1}{2}$ and $\alpha = \frac{1}{3}$ pretty well cover the expected range.

Values of V and of the upward reflection coefficient (for energy) $R_u = |V|^2$ have been obtained by expressing the Hankel functions in equation (15) in terms of Bessel functions of the first kind, and using tabled values of the Bessel functions of fractional order. Table 1 gives values of R_u for different dimensionless wavenumbers a. For $a \gg 1$, corresponding to the short wavelengths (15–40 km) discussed in the previous section, we can find a simple asymptotic expression for R_u. Using Hankel's asymptotic expansion

$$H_{\nu}^{(2)}(\eta) \approx \frac{2}{\pi\eta}\left[1 - i\left(\frac{4\nu^2 - 1}{8\eta}\right) + \cdots\right] \times \exp\left\{-i\left[\eta - \left(\frac{\nu}{2} + \frac{1}{4}\right)\pi\right]\right\}, \quad (\eta \gg 1),$$

we find, to first order in a^{-1},

$$V = \frac{i(1 - \alpha)}{4a},$$

and hence

$$R_u = |V|^2 = \frac{(1 - \alpha)^2}{16a^2}, \quad a \gg 1. \tag{17}$$

TABLE 1

VALUES OF THE UPWARD REFLECTION COEFFICIENT

$\frac{a}{\alpha+1}$	$\alpha = \frac{1}{2}$		$\alpha = \frac{1}{3}$	
	λ/H_0	R_u	λ/H_0	R_u
10.....	8.38×10^{-1}	6.9×10^{-5}	9.42×10^{-1}	1.6×10^{-4}
5.....	1.68	2.7×10^{-4}	1.88	6.3×10^{-4}
1.....	8.38	5.0×10^{-3}	9.42	1.1×10^{-2}
5×10^{-1}.....	1.68×10^{1}	1.4×10^{-2}	1.88×10^{1}	3.2×10^{-2}
1×10^{-1}.....	8.38×10^{1}	8.0×10^{-2}	9.42×10^{1}	1.7×10^{-1}
5×10^{-2}.....	1.68×10^{2}	1.4×10^{-1}	1.88×10^{2}	2.8×10^{-1}
1×10^{-2}.....	8.38×10^{2}	3.1×10^{-1}	9.42×10^{2}	5.6×10^{-1}
5×10^{-3}.....	1.68×10^{3}	4.0×10^{-1}	1.88×10^{3}	6.6×10^{-1}
1×10^{-3}.....	8.38×10^{3}	5.8×10^{-1}	9.42×10^{3}	8.3×10^{-1}

This expression is accurate to two significant figures for $a \geq 5$.

Table 1 and expression (17) show that the upward reflection from the umbral subphotosphere is very weak. For $\lambda = 40$ km ($a = 10\pi$) and $\alpha = \frac{1}{2}$, equation (17) gives $R_u = 1.58 \times 10^{-5}$. Of course, we cannot put any observational limit on the wavelength of Alfvén waves that might be propagating downward into the Sun. But even if these waves had wavelengths comparable to the size of the whole umbra, the reflection coefficient is still small ($R_u = 0.175$ for $\alpha = \frac{1}{3}$, $\lambda = 94.25H_0 = 9425$ km).

V. DISCUSSION

The model umbra considered here predicts total downward reflection and negligible upward reflection of Alfvén waves in a sunspot. The downward reflection in the photosphere and chromosphere is so strong that very little wave energy reaches as high as the transition region and corona. Thus, if Alfvén waves are indeed cooling a sunspot, we conclude that these waves must propagate downward into the solar interior. This same conclusion was reached by Wilson (1975) and by Beckers and Schneeberger (1977) on the basis of observations.

The very small upward reflection coefficient for Alfvén waves in our model umbra indicates that cooling by downward-propagating waves is possible. However, because of great uncertainties about the structure of a sunspot below the solar surface, this result can be considered only as suggestive. If the spot is fairly deep and the magnetic flux rope nearly vertical in the convection zone, as suggested by Meyer *et al.* (1974), then the present model is fairly accurate. In order for the Alfvén waves to cool the sunspot, they must propagate sufficiently far from the surface layers of the sunspot before their energy is dissipated into heat and thoroughly mixed to become part of the overall radial heat flux in the solar interior. This process could be limited by scattering and dissipation of the Alfvén waves due to inhomogeneities in the density and magnetic field in the subsurface structure of the sunspot. Much more work needs to be done to fairly evaluate the possibility of cooling by downward-propagating Alfvén waves.

It should also be noted that the pure Alfvén waves considered here are a very special type of wave, involving only horizontal motions. Umbral convection may not produce this kind of motion very efficiently, and one should perhaps consider more general magneto-atmospheric waves with vertical motions and compression as well. But these more general waves will no doubt suffer much greater upward reflection than the pure Alfvén waves because of the increasing sound speed and rapidly changing buoyancy force in the umbral subphotosphere. This presents an additional difficulty for the wave-cooling hypothesis.

I am indebted to Jacques Beckers, Alfred Clark, Jr., and John Molyneux for helpful discussions. Most of this work was done during a visit at Sacramento Peak Observatory. This work was supported by the Air Force Geophysics Laboratory under contract F19628-77-C-0079.

APPENDIX

Here we present the calculation of the downward reflection coefficient R_d in the case where layer 3 (the corona) of the model umbra is included. A similar analysis has been given by Hollweg (1972) in a different context. The solution of the wave equation (2) in layer 3 that represents an outgoing wave as $z \to +\infty$ is

$$\hat{u} = CJ_0(\xi), \tag{A1}$$

where

$$\xi = b \exp\left[-(z - z_t)/2H_1\right], \qquad b = \frac{2H_1\omega}{v_{A_0}} s, \qquad s = \left(\frac{T_0}{T_1}\right)^{1/2} \exp\left(-z_t/2H_0\right), \tag{A2}$$

and $H_1 = RT_1/g$ is the scale height in layer 3. We have solution (7) in layer 1, solution (9) in layer 2 (with $B \neq 0$), and solution (A1) in layer 3. Matching $\hat{u}$ and $d\hat{u}/dz$ across $z = 0$ and $z = z_t$, we obtain (after some algebraic manipulation)

$$A = \frac{2[sJ_1(b)Y_0(\zeta_t) - J_0(b)Y_1(\zeta_t)]}{M + iN}, \tag{A3}$$

$$B = \frac{2[J_0(b)J_1(\zeta_t) - sJ_1(b)J_0(\zeta_t)]}{M + iN}, \tag{A4}$$

$$C = \frac{2[Y_0(\zeta_t)J_1(\zeta_t) - J_0(\zeta_t)Y_1(\zeta_t)]}{M + iN}, \tag{A5}$$

and

$$U = \frac{M - iN}{M + iN}, \tag{A6}$$

where $\zeta_t = a \exp(-z_t/2H_0)$ and

$$M = \{J_0(b)[J_1(\zeta_t)Y_0(a) - Y_1(\zeta_t)J_0(a)] - sJ_1(b)[J_0(\zeta_t)Y_0(a) - Y_0(\zeta_t)J_0(a)]\}, \tag{A7}$$

$$N = \{J_0(b)[J_1(\zeta_t)Y_1(a) - Y_1(\zeta_t)J_1(a)] - sJ_1(b)[J_0(\zeta_t)Y_1(a) - Y_0(\zeta_t)J_1(a)]\}. \tag{A8}$$

Since M and N are real, we have $R_d = |U|^2 = 1$ for all values of the parameters a, b, and ζ_t.

Considerable simplification of the coefficients A, B, and C is possible for our purposes. If we adopt the values $T_0 = 4500$ K, $T_1 = 10^6$ K, $H_0 = 100$ km, $z_t = 2000$ km, then $s = 3 \times 10^{-6}$, and it is sufficient to take $s \to 0$ in equations (A3)–(A8). The parameter ζ_t is also small as long as a is not too large ($\zeta_t \leq 5.7 \times 10^{-3}$ for $\lambda \geq 10$ km), and we can also take $\zeta_t \to 0$ in equations (A3)–(A8), giving

$$A = \frac{2}{J_0(a) + iJ_1(a)}, \qquad B = 0, \qquad C = \frac{2}{J_0(b)[J_0(a) + iJ_1(a)]}. \tag{A9}$$

In this limit the solution in layer 2 is the same as that obtained in § III where we ignored layer 3. Also, for $\lambda \geq 10$ km, $b \leq 0.085$ and $J_0(b) \approx 1$. Thus, the only important effect of including layer 3 for the parameter values assumed here is to change the length scale in the solution above $z = z_t$.

REFERENCES

Adam, J. A. 1976, *J. Phys. A*, **9**, L193.
Beckers, J. M. 1976, *Ap. J.*, **203**, 739.
Beckers, J. M., and Schneeberger, T. J. 1977, *Ap. J.*, **215**, 356.
Biermann, L. 1941, *Vierteljahrsschr. Astr. Gesellsch.*, **76**, 194.
Boruta, N. 1977, *Ap. J.*, **215**, 364.
Brekhovskikh, L. M. 1960, *Waves in Layered Media* (New York: Academic Press).
Clark, A., Jr. 1978, in preparation.
Cowling, T. G. 1976, *M.N.R.A.S.*, **177**, 409.
Danielson, R. E. 1965, in *IAU Symposium No. 22, Stellar and Solar Magnetic Fields*, ed. R. Lüst (Amsterdam: North-Holland), p. 315.
Eschrich, K.-O., and Krause, F. 1977, *Astr. Nach.*, **298**, 1.
Ferraro, V. C. A. 1954, *Ap. J.*, **119**, 393.
Ferraro, V. C. A., and Plumpton, C. 1958, *Ap. J.*, **127**, 459.
Geronicolas, E. A. 1977, *Ap. J.*, **211**, 966.
Hollweg, J. V. 1972, *Cosmic Electrodyn.*, **2**, 423.
Meyer, F., Schmidt, H. U., and Weiss, N. O. 1977, *M.N.R.A.S.*, **179**, 741.
Meyer, F., Schmidt, H. U., Weiss, N. O., and Wilson, P. R. 1974, *M.N.R.A.S.*, **169**, 35.
Moore, R. L. 1973, *Solar Phys.*, **30**, 403.
Musman, S. 1967, *Ap. J.*, **149**, 201.
Parker, E. N. 1974*a*, *Solar Phys.*, **36**, 249.
———. 1974*b*, *Solar Phys.*, **37**, 127.
———. 1975*a*, *Solar Phys.*, **40**, 275.
———. 1975*b*, *Solar Phys.*, **40**, 291.
———. 1977, *M.N.R.A.S.*, **179**, 93p.
Roberts, B. 1976, *Ap. J.*, **204**, 268.
Savage, B. D. 1969, *Ap. J.*, **156**, 707.
Spruit, H. C. 1974, *Solar Phys.*, **34**, 277.
———. 1977, *Solar Phys.*, **55**, 3.
Uchida, Y., and Kaburaki, O. 1974, *Solar Phys.*, **35**, 451.
Uchida, Y., and Sakurai, T. 1975, *Pub. Astr. Soc. Japan*, **27**, 259.
Watson, G. N. 1966, *A Treatise on the Theory of Bessel Functions* (2d ed.; Cambridge: Cambridge University Press), p. 97.
Wilson, P. R. 1971, *Solar Phys.*, **21**, 101.
———. 1975, *Solar Phys.*, **42**, 333.

JOHN H. THOMAS: Department of Mechanical and Aerospace Sciences, University of Rochester, Rochester, NY 14627

THERMAL MODELS OF SUNSPOTS

ALFRED CLARK, JR.

Department of Mechanical and Aerospace Sciences and C. E. Kenneth Mees Observatory, University of Rochester, Rochester, N.Y. 14627, U.S.A.

(Received 20 July; in revised form 3 December, 1978)

Abstract. We study simple thermal models of sunspots based on the concept of partial inhibition of convection by strong magnetic fields. As in other similar studies (Parker, 1974a; Eschrich and Krause, 1977; Spruit, 1977a), the calculations involve solutions of the heat equation with an appropriate distribution of thermal conductivity.

The simplicity of the present model allows a detailed study of the dependence of the solution on the spot parameters, such as the depth of the region in which convection is inhibited. The most important specific results from the model are: (1) the edge of the umbra is sharp, even for deep spots; (2) deep spots produce weak bright rings in the surrounding atmosphere, whereas shallow spots produce intense rings which are difficult to reconcile with observations; (3) only a surface layer of a spot, with thickness of the order of the temperature scale height, is cool.

The present model, like those of Eschrich and Krause (1977) and Spruit (1977a) yields surface temperature distributions resembling sunspots. Since the three models all use different descriptions of the convective heat transport, we conclude that the major predictions of the inhibition theory are relatively insensitive to model details.

1. Introduction

Many years ago, Biermann (1941) suggested that a sunspot is dark because the strong magnetic field of the spot suppresses the convection that normally transports the solar flux. Discussions by Hoyle (1949) and Cowling (1953) called attention to the possibility that the magnetic field reduces, but does not suppress entirely, the convective heat flux. Later magnetohydrostatic models of sunspots by Chitre (1963), Deinzer (1965), Chitre and Shaviv (1967), Yun (1970), Mullan (1973), and Busse (1973) incorporated this concept in various ways.

Interest in the problem has been revived by Parker's (1974a, b, 1975a, b) vigorous advocacy of an alternative mechanism, originally suggested by Danielson (1965), whereby sunspots are cooled by the radiation of Alfvén waves. A number of recent papers, in addition to Parker's, have presented calculations bearing on the Alfvén wave mechanism (e.g., Roberts, 1976; Boruta, 1977; Thomas, 1978), or the inhibition of convection. The subject is controversial (Cowling, 1976a, Parker, 1977) and most authors have dealt with one or the other of the two mechanisms (the present work is no exception). However, it is worth emphasizing that they are not mutually exclusive. It may well be that both are important in sunspots. The work reported here is concerned entirely with the mechanism of inhibition of convection, so, apart from an occasional comment, we do not discuss further the cooling by Alfvén waves.

The model presented here is related to the models of Parker (1974a), Eschrich and Krause (1977), and Spruit (1977a). All four models are appropriately called thermal

Solar Physics **62** (1979) 305–330. 0038–0938/79/0622–0305 $03.90.

models, because they involve the calculation of a temperature distribution from some form of heat transport equation. These models do not deal with the mechanical aspects of a sunspot. Thus they provide a necessary but not sufficient test for a proposed cooling mechanism. The additional constraints imposed by mechanical equilibrium may rule out mechanisms which are allowed by thermal models. Nevertheless, the models are very useful in studying the thermal balance of a sunspot.

The four thermal models under present discussion are all based on the assumption that the turbulent convective heat transport can be modeled by some kind of conductive law. There is no unique, or even generally accepted way to do this, and, in fact, such an assumption is, with present knowledge, unverifiable. Because of this, one must accept at the outset the impossibility of absolutely verifying or rejecting the inhibition concept on the basis of thermal models. The proper objective is much more modest: namely, an attempt to assess plausibility. If a number of different ways of describing the turbulent convective heat transport lead to models which look like sunspots, then the magnetic inhibition concept becomes more plausible. If, on the other hand, only very special circumstances can yield a model resembling a spot, then some doubt is cast on the role of magnetic inhibition.

Consider now in more detail the nature of the four thermal models. In each case, one solves the steady-state heat equation in a half-space, with the heat flux asymptotically equal to the solar flux deep below the surface, and with an appropriate boundary condition (usually the Stefan-Boltzmann law) on the upper surface. The success of the model is then determined by comparing the calculated distribution of surface temperature with temperature distributions observed in sunspots. In constructing a model, there are three major choices: (i) the geometry of the region of strong magnetic field where the convective heat transport is affected (for brevity, we will call this region the spot in what follows); (ii) the form of the conductive law outside the spot; (iii) the alteration of the conductive law within the spot. We now summarize briefly how these three choices were made in the four models under discussion.

Eschrich and Krause (1977) chose a right circular cylinder for the spot shape. Outside of the spot, they used a Fourier law for the heat flux with an isotropic, constant conductivity. Within the spot they used a constant anisotropic conductivity, with the vertical conductivity (along the magnetic field) being considerably less than the horizontal conductivity.

Spruit (1977a) also used a cylindrical spot geometry. He used a more complicated description of the heat transport, based on mixing length theory. His conductivity was both depth-dependent and anisotropic. Outside the spot, the anisotropy was small. Inside the spot, Spruit took a vanishing horizontal conductivity. Thus although Eschrich and Krause (1977) and Spruit (1977a) agree that the convective transport is anisotropic, they do not agree on the sense of the anisotropy. As we discussed earlier, this kind of ambiguity is inherent in any attempt to model turbulent convective transport.

The present work fits into the catalog of models between the two works discussed above, since an isotropic conductivity is used both inside and outside the spot. In this respect, our model is like Parker's (1974a). Parker carried out calculations only for very shallow spots, however, whereas the present work has been done for a full range of aspect ratios.

In addition to filling a gap in a catalog, the work presented here has a second objective: to develop the simplest possible model of the inhibition concept. In the absence of an accurate theory of convective heat transport, one can make a strong case for simplicity. The principal advantage of a simple model is that it allows a thorough discussion of the dependence of the solution on the basic spot parameters.

The mathematical formulation for the present model is given in Section 2. In Section 3 some two-dimensional models are analyzed in detail. The analysis leads to some simple asymptotic approximations which are used to advantage in the discussion of three-dimensional models (Section 4). The reader not interested in mathematical details is invited to skip to Section 5, where conclusions from this work and other thermal models are discussed. In order to make Section 5 reasonably complete, the discussions of the solution, which would normally be scattered through Sections 2, 3, and 4, are deferred until Section 5.

2. Formulation

The geometry of the model is shown in Figure 1. We assume that the heat flux $\mathbf{Q}$ and the superadiabatic temperature gradient are linearly related. The thermal conductivity is assumed isotropic, with a reduced value in the volume V_1 where the inhibition is effective. Thus we take

$$\mathbf{Q} = \begin{cases} -K(\nabla T - \Gamma \mathbf{k}) & \text{in} \quad V_2 \\ -K(1-\varepsilon)(\nabla T - \Gamma \mathbf{k}) & \text{in} \quad V_1 , \end{cases} \tag{1}$$

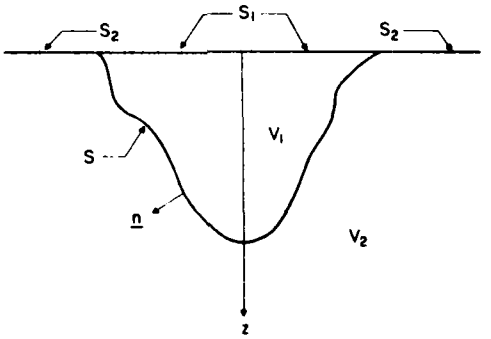


Fig. 1. Model geometry. The volume V_1 is the region of low thermal conductivity, and V_2 is the region of high conductivity, with S the interface between them. The gas radiates into the half-space $z < 0$ from the surface $z = 0$, which is divided into S_1 (bounding V_1) and S_2 (bounding V_2). The heat flux is asymptotically equal to $F_\odot$ as $z \to +\infty$.

where T is the temperature, Γ is the adiabatic gradient (assumed constant), $\mathbf{k}$ is a unit vector in the z-direction, K is the thermal conductivity outside the spot, and ε is the fractional reduction in conductivity within the spot. Both K and ε are taken to be spatially constant. The steady-state heat equation is $\nabla \cdot \mathbf{Q} = 0$, and this becomes

$$\nabla^2 T = 0 \quad \text{in } V_1 \text{ and } V_2 . \tag{2}$$

The temperature and the normal component of heat flux must be continuous at the interface S and V_1 and V_2. Thus

$$[T]_1^2 = 0$$

and

$$\mathbf{n} \cdot (\nabla T - \Gamma \mathbf{k})|_2 = (1-\varepsilon)\mathbf{n} \cdot (\nabla T - \Gamma \mathbf{k})|_1 \quad \text{on} \quad S. \tag{3}$$

At the upper surface, the gas radiates according to the Stefan-Boltzmann law, so we have

$$K(1-\varepsilon)\left(\frac{\partial T}{\partial z} - \Gamma\right) = \sigma T^4 \quad \text{on} \quad S_1$$

and

$$K\left(\frac{\partial T}{\partial z} - \Gamma\right) = \sigma T^4 \quad \text{on} \quad S_2 , \tag{4}$$

where σ is the Stefan–Boltzmann constant. The final condition is the asymptotic one far beneath the spot, where the heat flux must approach the solar flux $F_\odot$. Thus

$$\frac{\partial T}{\partial z} \underset{z \to \infty}{\sim} \Gamma + \frac{F_\odot}{K} . \tag{5}$$

(Here our approach differs from that of Eschrich and Krause (1977), whose imposed exact flux uniformity at the base of their cylindrical volume V_1.)

In the absence of a spot ($\varepsilon = 0$) the temperature gradient is constant, and the flux and surface temperature T_{S0} are related by

$$F_\odot = \sigma T_{S0}^4 . \tag{6}$$

The temperature distribution is then

$$T_0(z) = T_{S0}(1 + z/h) , \tag{7}$$

where

$$h = \left(\frac{1}{T}\frac{\mathrm{d}T}{\mathrm{d}z}\right)^{-1}_{z=0} = \left(\frac{\Gamma}{T_{S0}} + \frac{\sigma T_{S0}^3}{K}\right)^{-1} \tag{8}$$

is the temperature scale height at the surface. Another important scale, the superadiabatic temperature scale height at the surface, is defined by

$$H = \left(\frac{\mathrm{d}T/\mathrm{d}z - \Gamma}{T}\right)^{-1}_{z=0} = \frac{K}{\sigma T_{S0}^3} . \tag{9}$$

Returning to the general case, we introduce T_{S0} as a temperature scale and L as a length scale (to be chosen later, but always of the order of the spot diameter). Also, we separate out the behavior of T at $z = \infty$. Thus we introduce

$$\hat{z} = L^{-1}z, \qquad \hat{\nabla} = L\nabla,$$

and a new dependent variable Φ, defined by

$$T = T_{S0}\left(1 + \frac{L}{h}\hat{z} + \frac{\lambda\varepsilon}{4}\Phi\right), \tag{10}$$

where

$$\lambda = (4L/H) \tag{11}$$

is of the order of twice the spot diameter divided by the superadiabatic temperature scale height. With this scaling, the problem becomes:

$$\hat{\nabla}^2\Phi = 0 \quad \text{in } V_1 \text{ and } V_2, \tag{12}$$

$$\Phi \to 0 \quad \text{as} \quad \hat{z} \to \infty, \tag{13}$$

$$[\Phi]_1^2 = 0 \quad \text{on } S, \tag{14}$$

and

$$\mathbf{n}\cdot\hat{\nabla}\Phi|_2 = (1-\varepsilon)\mathbf{n}\cdot\hat{\nabla}\Phi|_1 - \mathbf{n}\cdot\mathbf{k} \quad \text{on} \quad S. \tag{15}$$

The upper surface boundary condition is now given by

$$\frac{\partial\Phi}{\partial\hat{z}} = \begin{cases} \dfrac{(1+\lambda\varepsilon\Phi/4)^4 - 1 + \varepsilon}{\varepsilon(1-\varepsilon)} & \text{on} \quad S_1 \\[2ex] \dfrac{(1+\lambda\varepsilon\Phi/4)^4 - 1}{\varepsilon} & \text{on} \quad S_2. \end{cases} \tag{16}$$

Because of the boundary condition (16), the problem is nonlinear. The solution would require elaborate numerical work, which can hardly be justified in view of the other limitations of the model. Thus we linearize the problem, based on the assumption of small ε. This restricts our ability to make quantitative statements about temperature amplitudes, but we can still have some confidence in the spatial scales emerging from the calculation. In Section 5, we discuss a crude but not unreasonable way to restore some of the nonlinearity omitted from the calculations. (Some comments on other models are in order here. Parker (1974a) also linearized the radiation condition. In his case, the small parameter was the ratio of the spot depth to the temperature scale height. In the present work, we have preferred to take the conductivity contrast ε to be small, since we wish to study deep spots. Eschrich and Krause (1977) also linearized the radiation condition. From the point of view of mathematical consistency, their linearization cannot be justified, since neither the spot depth nor the conductivity contrast is small in their calculations. However, as

they point out, the error associated with their linearization is probably less than 25%.)

The linearization is carried out by expanding Φ in powers of ε and keeping only the first term in the expansion. The result is:

$$\hat{\nabla}^2\Phi = 0 \quad \text{in } V_1 \text{ and } V_2, \tag{17}$$

$$\Phi \to 0 \quad \text{as} \quad \hat{z} \to \infty, \tag{18}$$

$$[\Phi]_1^2 = 0, \quad \mathbf{n}\cdot[\hat{\nabla}\Phi]_1^2 = -\mathbf{n}\cdot\mathbf{k} \quad \text{on} \quad S, \tag{19}$$

and

$$\frac{\partial\Phi}{\partial\hat{z}} - \lambda\Phi = \begin{cases} 1 & \text{on} \quad S_1 \\ 0 & \text{on} \quad S_2. \end{cases} \tag{20}$$

In this formulation, it becomes clear that the character of the solution depends only on the shape of the spot and the parameter λ. As we shall see later, the case of interest is $\lambda \gg 1$.

Although the problem defined by (17)–(20) is linear, it is made somewhat difficult by the jump conditions on the internal interface S. The basic technique we use is first to find a harmonic function χ which satisfies the jump condition (19) and the condition (18) at ∞, but not, in general, the surface boundary condition (20). Thus we let χ be a function which satisfies

$$\hat{\nabla}^2\chi = 0 \quad \text{in } V_1 \text{ and } V_2, \quad \chi \underset{z\to\infty}{\to} 0, \tag{21}$$

$$[\chi]_1^2 = 0, \qquad \mathbf{n}\cdot[\hat{\nabla}\chi]_1^2 = -\mathbf{n}\cdot\mathbf{k} \quad \text{on} \quad S. \tag{22}$$

To make χ unique, we also specify that

$$\chi = 0 \quad \text{on} \quad S_1 + S_2. \tag{23}$$

Once we have found such a χ, we define a new dependent variable Ψ by

$$\Phi = \chi + \Psi. \tag{24}$$

Then Ψ is a solution of the following boundary value problem:

$$\hat{\nabla}^2\Psi = 0 \quad \text{in} \quad V_1 + V_2, \tag{25}$$

$$\Psi \to 0 \quad \text{as} \quad \hat{z} \to \infty, \tag{26}$$

and

$$\frac{\partial\Psi}{\partial\hat{z}} - \lambda\Psi = f \quad \text{on} \quad \hat{z} = 0, \tag{27}$$

where

$$f = \begin{cases} -\dfrac{\partial\chi}{\partial z} + 1 & \text{on} \quad S_1 \\ -\dfrac{\partial\chi}{\partial z} & \text{on} \quad S_2. \end{cases} \tag{28}$$

The problem (25)–(28) involves no internal interfaces and is an ordinary half-space problem which can be solved by the Fourier transform. Thus we can deal with any spot geometry for which a 'jump' function χ can be found. As we shall see, this is easily done for elliptical spots in the two-dimensional case, and both prolate and oblate spheroidal spots in the three-dimensional case.

3. Two-Dimensional Models

3.1. Solution of Half-Space Problem

In this section, we solve the problem (25)–(28) for $\Psi(x, z)$, where x is the horizontal coordinate. We choose the length scale L so that the spot volume V_1 intersects the surface $z = 0$ along $-1 \leq x \leq 1$. The dimensional spot width is then $2L$. The problem is easily solved by a Fourier transform. With the aid of the convolution theorem, one can put the solution in the form

$$\Psi(x, z) = \int_{-\infty}^{\infty} G(x - x', z) f(x')\, dx', \tag{29}$$

where

$$G(x, z) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-(ikx + |k|z)}}{|k| + \lambda}\, dk. \tag{30}$$

It is easy to show (with the formulas in Chapter 5 of Abramowitz and Stegun, 1964) that the kernel $G(x, z)$ can be expressed in terms of an exponential integral E_1:

$$G(x, z) = -\mathrm{Re}\left\{\frac{e^{\xi}}{\pi} E_1(\xi)\right\}, \tag{31}$$

where

$$\xi = \lambda(z + ix), \tag{32}$$

and

$$E_1(\xi) = \int_{\epsilon}^{\infty} \frac{e^{-t}}{t}\, dt. \tag{33}$$

An important special case is $G(x, 0)$, which is needed in the calculation of the surface temperature. By starting from (30) with $z = 0$, one can show that

$$G(x, 0) = -\frac{1}{\pi} g(\lambda |x|), \tag{34}$$

where g is an auxiliary function of the exponential integrals, defined by (Abramowitz

and Stegun, 1964; p. 232)

$$g(u) = -\cos(u)Ci(u) - \sin(u)si(u)\,, \tag{35}$$

where

$$Ci(u) = \delta + \ln(u) + \int_0^u \frac{\cos t - 1}{t}\,dt \tag{36}$$

and

$$si(u) = -\frac{\pi}{2} + \int_0^u \frac{\sin t}{t}\,dt\,, \tag{37}$$

where $\delta = 0.577\,215\,664\,9$ is Euler's constant. For small u, the most useful representation of g is obtained from the expansions

$$Ci(u) = \delta + \ln(u) + \sum_{n=1}^{\infty} \frac{(-1)^n u^{2n}}{2n(2n)!} \tag{38}$$

and

$$si(u) = -\frac{\pi}{2} + \sum_{n=0}^{\infty} \frac{(-1)^n u^{2n+1}}{(2n+1)(2n+1)!}\,. \tag{39}$$

For $1 \le u < \infty$, the most useful representation is the rational approximation (Hastings, 1955)

$$g(z) = \frac{1}{u^2}\left[\frac{u^8 + a_1u^6 + a_2u^4 + a_3u^2 + a_4}{u^8 + b_1u^6 + b_2u^4 + b_3u^2 + b_4}\right] + \varepsilon(u)\,, \tag{40}$$

where

$$\begin{aligned} a_1 &= 42.242\,855, & b_1 &= 48.196\,927, \\ a_2 &= 302.757\,865, & b_2 &= 482.485\,984, \\ a_3 &= 352.018\,498, & b_3 &= 1114.978\,885, \\ a_4 &= 21.821\,899, & b_4 &= 449.690\,326, \end{aligned} \tag{41}$$

with $|\varepsilon(u)| < 3 \times 10^{-7}$ for $1 \le u < \infty$.

The parameter $\lambda = 4L/H$ is important in the solution, so we digress briefly to discuss typical values. From the model of the convection zone given by Spruit (1977b), we estimate that $H \sim 500$ km at $\tau = 1$. The diameter of the region of inhibition of convection in our model is $2L$, and it is logical to identify this with a typical umbral diameter – say $2L \sim 20\,000$ km. This gives $\lambda \sim 80$. There are uncertainties. In particular, H may well be different in other convection zone models, since it is determined by the superadiabatic gradient in the uppermost layers of the convection zone – a particularly difficult quantity to predict. Observed umbral

diameters vary widely, so L is also uncertain. In spite of the uncertainties, it is clear that λ is large, and that is all that is necessary for the calculations here. When a numerical value is required, we have taken $\lambda = 100$ for purposes of illustration.

It is worth noting that

$$\frac{1}{\pi}\int_{-\infty}^{\infty} g(|u|)\,du = 1\,, \tag{42}$$

a result that is most easily proved from the Fourier transform of $G(x, 0)$. Thus for large λ, we expect

$$\lim_{\lambda\to\infty}\frac{\lambda}{\pi}g(\lambda|x-x'|) = \delta(x-x')\,. \tag{43}$$

The formulas given here reduce the problem for Ψ to quadrature. To go further, we must choose a geometry for the spot and find the corresponding jump function χ.

3.2. Elliptical spots

We seek a function $\chi(x, z)$ satisfying (21)–(23), when the interface S is given by

$$x^2 + (z/D)^2 = 1\,, \tag{44}$$

where D is the aspect ratio of the spot. The determination of a suitable χ is straightforward, although tedious, and can be carried out with the aid of formulas for the solution of Laplace's equation in elliptic coordinates (Morse and Feshbach, 1953, p. 1195–1200). We skip the lengthy derivation, and only quote the results. We begin with shallow spots ($D < 1$). We have

$$\chi(x, z) = \begin{cases} z/(D+1) & \text{in} \quad V_1 \\ [D/(1-D^2)^{1/2}]e^{-m}\sin\theta & \text{in} \quad V_2\,, \end{cases} \tag{45}$$

where

$$m = \cosh^{-1}\left\{\frac{[(x+\{1-D^2\}^{1/2})^2+z^2]^{1/2}+[(x-\{1-D^2\}^{1/2})^2+z^2]^{1/2}}{2\{1-D^2\}^{1/2}}\right\} \tag{46}$$

with $m \geq 0$, and where

$$\theta = \cos^{-1}\left\{\frac{[(x+\{1-D^2\}^{1/2})^2+z^2]^{1/2}-[(x-\{1-D^2\}^{1/2})^2+z^2]^{1/2}}{2\{1-D^2\}^{1/2}}\right\} \tag{47}$$

with $0 \leq \theta \leq \pi$. The inversion of (46) and (47) is also useful:

$$\begin{aligned} x &= (1-D^2)^{1/2}\cosh m\cos\theta\,, \\ z &= (1-D^2)^{1/2}\sinh m\sin\theta\,. \end{aligned} \tag{48}$$

For deep spots ($D > 1$) we have

$$\chi(x, z) = \begin{cases} z/(D+1) & \text{in} \quad V_1 \\ [D/(D^2-1)^{1/2}]e^{-\mu}\cos\phi & \text{in} \quad V_2, \end{cases} \tag{49}$$

where

$$\mu = \cosh^{-1}\left\{\frac{[(z+\{D^2-1\}^{1/2})^2+x^2]^{1/2}+[(z-\{D^2-1\}^{1/2})^2+x^2]^{1/2}}{2\{D^2-1\}^{1/2}}\right\} \tag{50}$$

with $\mu \geq 0$, and

$$\phi = \cos^{-1}\left\{\frac{[(z+\{D^2-1\}^{1/2})^2+x^2]^{1/2}-[(z-\{D^2-1\}^{1/2})^2+x^2]^{1/2}}{2\{D^2-1\}^{1/2}}\right\} \tag{51}$$

with $-\pi/2 \leq \phi \leq \pi/2$. The transformation inverse to (50) and (51) is

$$\begin{aligned} x &= (D^2-1)^{1/2}\sinh\mu\sin\phi, \\ z &= (D^2-1)^{1/2}\cosh\mu\cos\phi. \end{aligned} \tag{52}$$

For semi-circular spots ($D = 1$), we have

$$\chi(x, z) = \begin{cases} \frac{1}{2}z & \text{in} \quad V_1 \\ \frac{1}{2}\dfrac{z}{x^2+z^2} & \text{in} \quad V_2. \end{cases} \tag{53}$$

The boundary function $f(x)$, needed in the calculation of $\Psi(x, z)$, is defined by (28). For $D \neq 1$, f is given by

$$f(x) = \begin{cases} \dfrac{D}{D+1} & \text{for} \quad |x| < 1 \\ \dfrac{D}{D^2-1}\left(\dfrac{|x|}{[D^2-1+x^2]^{1/2}}-1\right) & \text{for} \quad |x| > 1. \end{cases} \tag{54}$$

For a semi-circular spot ($D = 1$), we have

$$f(x) = \begin{cases} \frac{1}{2} & \text{for} \quad |x| < 1 \\ -\dfrac{1}{2x^2} & \text{for} \quad |x| > 1. \end{cases} \tag{55}$$

With $\chi(x, z)$ and $f(x)$ known, we can now compute the surface temperature.

3.3. Surface temperature

We will show that for large λ, a very simple approximation gives useful information. This approximation proves invaluable in the analysis of three-dimensional spots in Section 4. In addition, our calculations will elucidate the nature of the umbral boundary, whose sharpness has been the object of much discussion in the literature.

We begin with the exact expression for $\Psi(x, 0)$, obtained from (29) and (34):

$$\Psi(x,0) = -\frac{1}{\pi}\int_{-\infty}^{\infty} g(\lambda|x-x'|)f(x')\,dx'. \tag{56}$$

The fact that λ is large suggests that we use the delta-function approximation (43) for g. This yields

$$\lambda\Psi(x,0) \approx -f(x). \tag{57}$$

This result is also strongly suggested by the boundary condition (27). The approximation (57) is very attractive, since $f(x)$ is known in closed form. However, it cannot be uniformly valid in x, since $f(x)$ is discontinuous at $x = \pm 1$, whereas $\Psi(x, 0)$ is continuous. This is the classic signal of a singular perturbation problem, and we must use stretched coordinates in the vicinity of $x = \pm 1$ in order to remedy the defects in (57) (see, e.g., Van Dyke, 1975, for a general discussion). Since $\Psi(x, 0)$ is even in x, we need consider only $x = 1$. There, as one can show, the proper stretched coordinate is

$$\eta = \lambda(x-1). \tag{58}$$

We introduce this and $\eta' = \lambda(x'-1)$ in (56). We then take the limit $\lambda \to \infty$ at fixed η. After some elementary manipulations in which we make use of (42), we can put the result in the form

$$\lim_{\substack{\lambda\to\infty \\ \eta \text{ fixed}}} \lambda\Psi(x,0) = \Omega(\eta) = -\tfrac{1}{2}(f_- + f_+) - \left(\frac{f_+ - f_-}{\pi}\right)\operatorname{sgn}(\eta)\int_0^{|\eta|} g(u)\,du, \tag{59}$$

where $f_\pm = f(1_\pm)$ are the right and left limits of f at $x = 1$. The inner solution $\Omega(\eta)$ blends smoothly with the outer solution $\lambda\Psi(x, 0)$, since

$$\lim_{\eta\to\pm\infty} \Omega(\eta) = -f_\pm, \tag{60}$$

a result easily established from (42) and (59). The smooth variation of $\lambda\Psi$ across the umbral boundary is given by (59). We see that $\lambda\Psi$ varies from f_- to f_+ for $\Delta\eta \sim 1$. Thus from (58) we conclude that the thickness of this transition region is of the order of λ^{-1}, where λ is given by (11). Thus *the horizontal width of the region of rapidly changing temperature near the umbral boundary is of the order of the superadiabatic temperature scale height.*

For computational purposes, it is highly desirable to blend (57) and (59) into a uniformly valid approximation (again see Van Dyke, 1975, for a general discussion). This is easily done and the result is

$$\lambda\Psi \approx \begin{cases} -f(x) + \Omega(\eta) + f_-, & \eta < 0 \\ -f(x) + \Omega(\eta) + f_+, & \eta > 0. \end{cases} \tag{61}$$

We now have three representations for $\Psi(x, 0)$: the simplest approximation (57), the more complete approximation (61), and the exact representation (56). From their derivation, we know that (57) and (61) are asymptotically valid for large λ. However, only by computation and comparison with the exact solution (56) can we determine whether the approximations are numerically useful for λ values of interest ($\lambda \sim 100$). For this reason we have carried out a detailed numerical comparison of (56), (57), and (61). The numerical integration of (56) involves four difficulties: (i) the infinite range of integration; (ii) the large parameter λ, requiring a fine grid; (iii) the logarithmic singularity of g at $x = x'$; (iv) the discontinuities in f at $x = \pm 1$. A detailed description of the numerical scheme actually used would take many pages. Since there are no major new features in the scheme, we forgo the description.

The relation between the surface temperature $T_S(x)$ and $\Psi(x, 0)$ is easily obtained from Equations (10), (23), and (24):

$$\frac{T_S(x) - T_{S0}}{\varepsilon T_{S0}} = \Delta T_S(x) = \frac{\lambda \Psi(x, 0)}{4}. \tag{62}$$

Here $\Delta T_S(x)$ is the fractional change in surface temperature per unit fractional

TABLE I

The change in surface temperature $\Delta T_S(x)$ for $\lambda = 100$ and $D = 5$

	Values of $\Delta T_S(x)$		
x	Exact, based on (56)	Approximate, based on (61)	Approximate, based on (57)
0.0	−0.2079	−0.2075	−0.2083
0.2	−0.2078	−0.2073	−0.2083
0.4	−0.2076	−0.2070	−0.2083
0.6	−0.2070	−0.2063	−0.2083
0.8	−0.2051	−0.2044	−0.2083
1.0	−0.0836	−0.0833	undefined
1.2	0.0362	0.0357	0.0397
1.4	0.0362	0.0358	0.0378
1.6	0.0350	0.0346	0.0359
1.8	0.0335	0.0331	0.0341
2.0	0.0320	0.0316	0.0324
2.2	0.0304	0.0301	0.0307
2.4	0.0289	0.0286	0.0292
2.6	0.0275	0.0272	0.0277
2.8	0.0261	0.0258	0.0262
3.0	0.0248	0.0245	0.0249
3.2	0.0235	0.0232	0.0236
3.4	0.0223	0.0221	0.0224
3.6	0.0212	0.0209	0.0212
3.8	0.0201	0.0199	0.0202
4.0	0.0191	0.0189	0.0191

TABLE II

The change in surface temperature $\Delta T_S(x)$ for x near 1, and for $\lambda = 100$ and $D = 5$

	Values of $\Delta T_S(x)$		
x	Exact, based on (56)	Approximate, based on (61)	Approximate, based on (57)
0.70	−0.2064	−0.2057	−0.2083
0.80	−0.2051	−0.2044	−0.2083
0.85	−0.2038	−0.2031	−0.2083
0.90	−0.2013	−0.2005	−0.2083
0.91	−0.2004	−0.1997	−0.2083
0.92	−0.1994	−0.1987	−0.2083
0.93	−0.1981	−0.1974	−0.2083
0.94	−0.1964	−0.1957	−0.2083
0.95	−0.1941	−0.1934	−0.2083
0.96	−0.1909	−0.1901	−0.2083
0.97	−0.1859	−0.1851	−0.2083
0.98	−0.1774	−0.1766	−0.2083
0.99	−0.1597	−0.1589	−0.2083
1.00	−0.0836	−0.0833	undefined
1.01	−0.0075	−0.0079	0.0416
1.02	0.0101	0.0097	0.0415
1.03	0.0185	0.0181	0.0414
1.04	0.0234	0.0230	0.0413
1.05	0.0266	0.0262	0.0412
1.06	0.0288	0.0284	0.0411
1.07	0.0304	0.0300	0.0410
1.08	0.0316	0.0312	0.0409
1.09	0.0326	0.0321	0.0408
1.10	0.0333	0.0329	0.0407
1.15	0.0354	0.0349	0.0402
1.20	0.0362	0.0357	0.0397
1.30	0.0365	0.0361	0.0387

change in thermal conductivity. Tables I and II give a numerical comparison of the two approximations and the exact solution for the parameter values $\lambda = 100$ and $D = 5$. It can be seen that the uniformly valid approximation (61) is in excellent agreement with the exact solution. Even the simple approximation (57) is very good except in the immediate vicinity of $x = 1$. Calculations for other parameter values ($\lambda = 50$ and 100; $D = 0.5$, 1.0, 2.0, 5.0, and 10.0) give similar results. Outside of the thermal transition region of the umbral boundary, the simple approximation (57) is quite adequate. We will take advantage of this in our discussion of three-dimensional spots in Section 4.

Because the sharp umbral edge is of some interest, the solution near $x = 1$ (both exact and approximate) is shown in Figure 2, again for the parameter values $\lambda = 100$, $D = 5$. As predicted by the asymptotic theory, the width of the region of sharp temperature change is of the order of λ^{-1}.

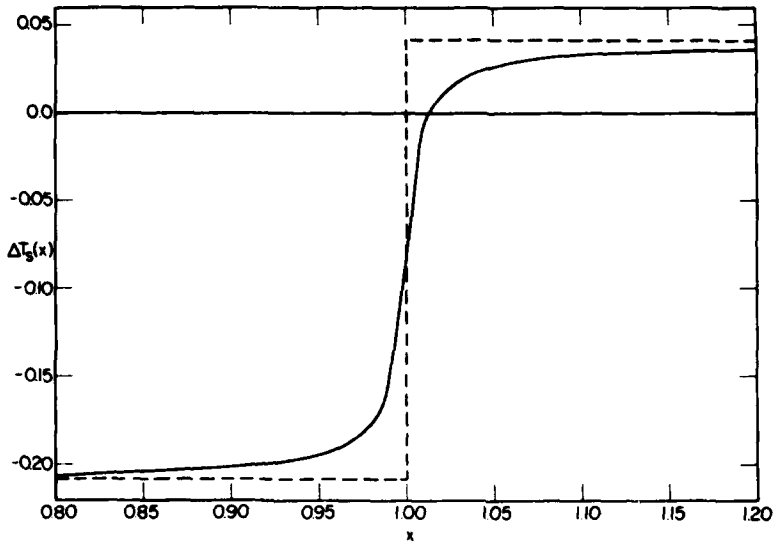


Fig. 2. Surface temperature distribution in the vicinity of the umbral edge ($x = 1$). The center of the umbra is at $x = 0$. The aspect ratio of the spot is $D = 5$. The value of λ is 100, corresponding to a spot diameter of 50 times the superadiabatic temperature scale height. The solid line is the exact solution (56) and the uniformly valid approximation (61), which are indistinguishable on this scale. The dashed line is the simple approximation (57).

3.4. Variation of Temperature with Depth

Although the surface temperature distribution is the principal observable quantity, the distribution of temperature below the visible surface is also of interest. In this section, we obtain some important qualitative results about the depth dependence of the temperature.

The starting point is Equation (10) for the temperature, Equation (24) for Φ, Equations (29)–(33) for Ψ, and Equations (45)–(55) for $\chi(x, z)$ and $f(x)$. The unperturbed temperature $T_0(z)$ is given by (7). A useful measure of the temperature distribution is $\Delta T(x, z)$, defined to be the fractional change in temperature per unit fractional change in conductivity:

$$\Delta T(x, z) = \frac{T(x, z) - T_0(z)}{\varepsilon T_0(z)} = \frac{\lambda(\chi + \Psi)}{4(1 + \lambda z)}. \tag{63}$$

For $z = 0$, this reduces to $\Delta T_S(x) = \lambda \Psi(x, 0)/4$, discussed in the previous section. Away from the surface, however, the term χ dominates. To show this, we use the expression (31) for G, and observe that, for $\lambda \gg 1$, and $\lambda z \gg 1$, the complex argument $\xi = \lambda[z + i(x - x')]$ is always large. Then we can use an asymptotic expansion for $E_1(\xi)$ (Abramowitz and Stegun, 1964; p. 231) to write

$$G \sim -\frac{1}{\pi} \operatorname{Re}\left\{\frac{1}{\xi} - \frac{1}{\xi^2} + \frac{2}{\xi^3} - \frac{3!}{\xi^4} + - \right\}, \qquad \xi \to \infty. \tag{64}$$

Substitution of this into Equation (29) for Ψ yields

$$\Psi(x, z) = -\frac{z}{\pi\lambda} \int_{-\infty}^{\infty} \frac{f(x')\,dx'}{(x-x')^2+z^2} + O(\lambda^{-2}) \quad \text{for} \quad \lambda z \gg 1. \tag{65}$$

Thus Ψ is small of order λ^{-1} provided $\lambda z \gg 1$. Since χ is of order one with respect to λ, we conclude that

$$\Delta T(x, z) \approx \frac{\chi}{4z} \quad \text{for} \quad \lambda z \gg 1. \tag{66}$$

Since $\chi > 0$ for $z > 0$, we get the important conclusion that the temperature perturbation is *positive* away from the upper boundary. Since Ψ is $O(\lambda^{-1})$ even on $z = 0$, we have $\Psi = O(\lambda^{-1})$ everywhere, and χ will dominate except where $\chi = 0$. It is easy to show that $\chi \leqslant O(\lambda^{-1})$ only for $\lambda z \leqslant 1$. *Thus the cooling of a spot is purely a surface phenomenon; only within one or two superadiabatic temperature scale heights of the visible surface is the spot cooler than the normal atmosphere. The remainder of the interior of the spot is hotter than the normal atmosphere.* A more quantitative discussion of this effect is given in Section 5.2 for three-dimensional spots.

4. Three-Dimensional Spots

We choose half spheroids for the geometry since they are sufficiently simple to allow an analytical solution, but sufficiently general to include deep and shallow spots. We choose our basic length scale L to be the radius of the horizontal circular section of the spheroid. The equation of the interface S is then

$$x^2 + y^2 + \frac{z^2}{D^2} = 1. \tag{67}$$

The discussion in Section 3 shows that all essential information can be obtained from the jump function χ by means of simple approximations. Thus we limit our calculations to the determination of χ. As in the two-dimensional case, the calculation is lengthy but straightforward. The necessary formulas are given by Morse and Feshbach (1953, pp. 1284–1294). As before, we skip the derivation and only quote the results. We begin with deep spots (prolate spheroids; $D > 1$). Then χ is most conveniently expressed in terms of

$$\xi(r, z) = \frac{\{[z + (D^2-1)^{1/2}]^2 + r^2\}^{1/2} + \{[z - (D^2-1)^{1/2}]^2 + r^2\}^{1/2}}{2(D^2-1)^{1/2}}, \tag{68}$$

where $r = (x^2 + y^2)^{1/2}$ is the cylindrical radial coordinate. The interface S is given by

$$\xi = \xi_0 = \frac{D}{(D^2-1)^{1/2}}, \tag{69}$$

and the spot volume V_1 corresponds to $1 \le \xi \le \xi_0$, with the exterior V_2 being $\xi_0 \le \xi < \infty$. Then $\chi(r, z)$ is given by

$$\chi = \begin{cases} \dfrac{Dz}{(D^2-1)^{3/2}}\left[\frac{1}{2}\ln\left(\dfrac{\xi_0+1}{\xi_0-1}\right)-\dfrac{1}{\xi_0}\right] & \text{in} \quad V_1 \\ \dfrac{Dz}{(D^2-1)^{3/2}}\left[\frac{1}{2}\ln\left(\dfrac{\xi+1}{\xi-1}\right)-\dfrac{1}{\xi}\right] & \text{in} \quad V_2 . \end{cases} \tag{70}$$

The boundary function f (defined by (28)) is

$$f(r) = \begin{cases} 1-\dfrac{D}{(D^2-1)^{3/2}}\left[\frac{1}{2}\ln\left(\dfrac{\xi_0+1}{\xi_0-1}\right)-\dfrac{1}{\xi_0}\right] & \text{for} \quad r<1 \\ -\dfrac{D}{(D^2-1)^{3/2}}\left[\frac{1}{2}\ln\left(\dfrac{\xi+1}{\xi-1}\right)-\dfrac{1}{\xi}\right] & \text{for} \quad r>1 , \end{cases} \tag{71}$$

where now

$$\xi = \xi(r, 0) = \left(1+\frac{r^2}{D^2-1}\right)^{1/2} . \tag{72}$$

Consider now shallow spots (oblate spheroids; $D<1$). The solution is most conveniently expressed in terms of

$$\rho = \frac{[r^2+z^2-1+D^2+\{(r^2+z^2-1+D^2)^2+4(1-D^2)z^2\}^{1/2}]^{1/2}}{[2(1-D^2)]^{1/2}} . \tag{73}$$

The interface S is given by

$$\rho = \rho_0 = \frac{D}{(1-D^2)^{1/2}} , \tag{74}$$

and the spot volume V_1 corresponds to $0 \le \rho \le \rho_0$, with the exterior V_2 being $\rho_0 \le \rho < \infty$. Then $\chi(r, z)$ is given by

$$\chi = \begin{cases} \dfrac{Dz}{(1-D^2)^{3/2}}\left[\dfrac{1}{\rho_0}-\tan^{-1}\left(\dfrac{1}{\rho_0}\right)\right] & \text{in} \quad V_1 \\ \dfrac{Dz}{(1-D^2)^{3/2}}\left[\dfrac{1}{\rho}-\tan^{-1}\left(\dfrac{1}{\rho}\right)\right] & \text{in} \quad V_2 . \end{cases} \tag{75}$$

Then the boundary function f is

$$f(r) = \begin{cases} 1-\dfrac{D}{(1-D^2)^{3/2}}\left[\dfrac{1}{\rho_0}-\tan^{-1}\left(\dfrac{1}{\rho_0}\right)\right] & \text{for} \quad r<1 \\ -\dfrac{D}{(1-D^2)^{3/2}}\left[\dfrac{1}{\rho}-\tan^{-1}\left(\dfrac{1}{\rho}\right)\right] & \text{for} \quad r>1 , \end{cases} \tag{76}$$

where now

$$\rho = \rho(r, 0) = \left(\frac{r^2}{1-D^2}-1\right)^{1/2} . \tag{77}$$

The special case of spherical spots ($D = 1$) can also be treated. One gets

$$\chi(r, z) = \begin{cases} \frac{1}{3}z & \text{in} \quad V_1 \\ \frac{1}{3}\frac{z}{(r^2+z^2)^{3/2}} & \text{in} \quad V_2, \end{cases} \tag{78}$$

and the boundary function f is

$$f(r) = \begin{cases} \frac{2}{3}, & r<1 \\ -\frac{1}{3r^3}, & r>1. \end{cases} \tag{79}$$

Consider now the surface temperature. We use the approximation (57) and the definition (62) to get

$$\Delta T_S(r) = \frac{T_S(r) - T_{S0}}{\varepsilon T_{S0}} = \frac{\lambda \Psi(r, 0)}{4} \approx -\frac{f(r)}{4}. \tag{80}$$

In this approximation, we do not resolve the sharp but smooth transition at the umbral boundary, but we do get good estimates of (i) the umbral temperature, (ii) the maximum temperature in the bright ring, and (iii) the horizontal scale for the fall-off of the temperature in the bright ring as we move away from the spot. In addition, this approximation satisfies the global conservation of energy. Let $\Delta F = 4\sigma T_{S0}^3 (T_S - T_{S0})$ be the (linearized) perturbation in the radiative flux at the free surface. Then the fractional flux perturbation is

$$\frac{\Delta F}{\sigma T_{S0}^4} = \frac{4(T_S - T_{S0})}{T_{S0}} = \varepsilon \lambda \Psi(r, 0) \approx -\varepsilon f. \tag{81}$$

Conservation of energy requires that the integral over $S_1 + S_2$ of ΔF vanish. It is straightforward to show this, by using the definition (28) for f, the jump conditions (22) for χ, the Equation (21) for χ, and the divergence theorem.

Finally, it is worth noting that within the framework of the large λ approximation used here, the character of the solution depends only on the aspect ratio D.

5. Discussion

The discussions based on the calculations of Sections 2, 3, and 4 all have been deferred to this section, in order to present them in a coherent and unified way. Specific points of interest are discussed in Sections 5.1 (the sharpness of the umbral boundary), 5.2 (the internal temperature of spots), 5.3 (bright rings and penumbrae), and 5.4 (spot depth). Concluding remarks and a summary are presented in 5.5.

5.1. The sharpness of the umbral boundary

Several authors (e.g., Cowling, 1953; Parker, 1974a) have suggested that the observed sharp umbral boundary is inconsistent with models based on the diffusion

equation, unless the depth of the region of thermal inhibition is very small. However, very shallow spots, in the inhibition theory, produce intense bright rings (Parker, 1974a) and are therefore inconsistent with observation (see Section 5.4 below for further discussion of spot depth). Cowling (1976a, b) has suggested that it is not necessary for the depth to be small, provided that the horizontal thermal conductivity in the region of strong field is greatly reduced. Even this constraint is not necessary, as the present model shows. Of course the horizontal conductivity in the spot *may* be small. The point is that the sharp umbral boundary does not require it. In fact, it is difficult not to get a sharp umbral boundary out of thermal models – the models of Eschrich and Krause (1977), Spruit (1977a) and the present work all show sharp umbral boundaries. We agree completely with Spruit (1977a) who says: "To maintain a sharp transition in temperature, a sharply bounded and strong inhomogeneity in the diffusion coefficient is sufficient." We would add to this that the conductivity inhomogeneity must be present at the visible surface, since it is really the combination of the inhomogeneity, the radiative boundary condition, and the small superadiabatic temperature scale height (compared to the spot size) which gives the sharp umbral boundary.

Neither Spruit nor Eschrich and Krause identify the thickness of the umbral temperature transition in terms of other scales. As our analysis in Section 3.3 shows, this horizontal transition region has a thickness equal to the superadiabatic temperature scale height at the surface (defined by (9)), provided that the spot dimensions greatly exceed this scale.

It should be emphasized that, even though thermal models predict the sharp umbral boundary, they do not explain it. What they show is that *if* there is a sharp change in the thermal conductivity, *then* there will be a sharp change in temperature at the visible surface. In the inhibition theory, a sharp change in thermal conductivity is caused by a sharp change in magnetic field strength. Interestingly enough, the theory of cooling by Alfvén waves, as developed by Parker (1974b), also requires a sharp change in magnetic field strength to get a sharp change in temperature. Thus while both theories can accommodate a sharp umbral boundary, there remains the fundamental question of why the magnetic flux tube of the spot has a sharp boundary.

5.2. The internal temperature of sunspots

As Parker (1974a) has shown for shallow spots, and as the present work (Section 3.4, Section 4) shows for both deep and shallow spots, the subsurface interior of a spot is hotter than the normal atmosphere at the same level. In fact, in a thermal model of a spot, the cool region extends downward only a distance of the order of the superadiabatic temperature scale height. It is of interest that considerations of mechanical equilibrium also suggest that sunspot cooling is confined to the surface layers (Weiss, 1964, 1969).

Consider now an estimate of the magnitude of this subsurface heating. We begin with the formula (10), from which we compute the dimensional vertical temperature

gradient in the spot volume (V_1):

$$\left.\frac{dT}{dz}\right|_{V_1} = \frac{T_{S0}}{H}\left(1+\varepsilon\frac{\partial\Phi}{\partial\hat{z}}\right). \tag{82}$$

At distances of several superadiabatic temperature scale heights or more below the visible surface, $\Phi \approx \chi$ (as discussed in Section 3.4). For the spheroidal spots of Section 4, the gradient in V_1 is constant, and we have

$$\frac{\partial\chi}{\partial\hat{z}} = \nabla_1 = \begin{cases} \dfrac{D}{(1-D^2)^{3/2}}\left[\dfrac{1}{C} - \tan^{-1}\left(\dfrac{1}{C}\right)\right] & \text{for } D<1 \\ \frac{1}{3} & \text{for } D=1 \\ \dfrac{D}{(D^2-1)^{3/2}}\left[\frac{1}{2}\ln\left(\dfrac{C+1}{C-1}\right) - \dfrac{1}{C}\right] & \text{for } D>1, \end{cases} \tag{83}$$

where

$$C = \frac{D}{|D^2-1|^{1/2}}. \tag{84}$$

Here the spot volume V_1 is one-half of a spheroid, and D is the ratio of the depth to the radius of the spheroid. To complete the estimate, we need a value of ε. We can relate ε to the umbral temperature T_u. From (10) evaluated at $z=0$, with the approximation (57) for Ψ and the definition (28) for f, we get

$$T_U = T_{S0}\left[1 - \frac{\varepsilon}{4}(1-\nabla_1)\right]. \tag{85}$$

We may solve (85) for ε. Then by substitution into (82) we get

$$\left.\frac{dT}{dz}\right|_{V_1} = \frac{T_{S0}}{H}(1+\alpha), \tag{86}$$

where

$$\alpha = 4(1 - T_U/T_{S0})\left(\frac{\nabla_1}{1-\nabla_1}\right). \tag{87}$$

Then α is the fractional excess temperature gradient in the spot. For illustrative purposes, we choose $T_U = 4000$ K and $T_{S0} = 5800$ K, and then calculate α as a function of the spot aspect ratio D. Some typical values of $\alpha(D)$ are $\alpha(0.5) = 1.384$, $\alpha(1.0) = 0.621$, $\alpha(2.0) = 0.261$, and $\alpha(5.0) = 0.073$. The numbers cannot be taken too seriously since the value of ε implied by (85) is not small, in violation of the mathematical basis of the calculation. Nevertheless the numbers suggest that the effect may be particularly significant for shallow spots. Since neither Spruit (1977a) nor Eschrich and Krause (1977) discuss the temperature variation beneath the surface, it is not known to what extent the present results are model-dependent.

Elementary physical reasoning about heat flow around obstacles suggests strongly that some subsurface temperature excess in the spot region is inevitable. With a conductivity which increases rapidly with depth, such as that used by Spruit (1977a) the excess temperature may be smaller than estimated here.

In the traditional version of the inhibition theory, one has usually assumed that the coolness of the spot extends over the spatial region in which inhibition is important. A very attractive feature of that assumption is that it leads to a mechanism for field concentration, first suggested by Parker (1955). In simplest terms, the mechanism is the following: increasing the field strength gives greater inhibition of convection, causing the spot to become cooler. As the spot cools, its pressure drops, and it must contract to preserve mechanical equilibrium. The contraction increases the field strength further, and the process continues until some kind of equilibrium is reached. (See Galloway *et al.* (1977) for a recent discussion of this and other field-concentration mechanisms.) As Parker (1974a, 1976) has emphasized, and as the present work shows, this mechanism cannot work in the inhibition theory, because the spot region is not cooled (except very near the visible surface) but is heated by the inhibition process. Parker (1974a) concludes that the heating would in fact disperse the field, and that this is a major objection to the inhibition theory. In the absence of a quantitative theory of spot formation, it is difficult to draw such a definite conclusion. We prefer the following more conservative conclusion: If the inhibition theory with increased subsurface temperatures is correct, then mechanical equilibrium (internal gas pressure plus magnetic pressure equals external gas pressure) requires that at any given level the spot have a lower density than the surrounding atmosphere. Thus some material must be expelled from a flux tube during the concentration phase of its history. Alternatively, as suggested by Meyer *et al.* (1974), the mechanical balance may be partially dynamic if there are significant large scale flows outside the flux tube. In any case, the above values of α suggest that the mechanical balance problem is more severe for shallow spots than deep ones.

A digression on cooling by Alfvén waves is useful here. In that theory, it is perfectly possible for the cooling to extend over a sizable volume, rather than being a surface effect as in the inhibition theory. Such a volume cooling would make sunspot formation much easier to understand. Cowling (1976a; see also Parker, 1977) has argued, on general thermodynamic grounds, that it is very difficult to transform all of the missing heat flux into Alfvén waves. However, that may not be necessary. As the calculations of this section show, deep spots, in the inhibition theory, have only slightly warm interiors. It is possible that a modest flux of Alfvén wave energy, in combination with the inhibition of convection, can produce a deep spot with a substantial cool volume. Such hybrid theories, as exemplified by the work of Mullan (1974), deserve more attention than they have received.

5.3. Bright rings and penumbrae

A point of primary interest in any sunspot theory is the disposition of the 'missing flux'. As many authors have pointed out, the inhibition theory requires a bright ring

around a spot, in which the diverted heat flux is radiated from the visible surface. The amplitude and scale of this bright ring are of great interest, since they are, in principle, observable. In this section, we will summarize the predictions of the present model, and compare them with those of Parker (1974a), Eschrich and Krause (1977), and Spruit (1977a). As we shall see, the comparison with observation is complex and inconclusive.

We begin with a discussion of the spatial scale of the bright ring. In calculating the surface temperature, we use the large λ approximation (57). Then we get from (10), with $z = 0$,

$$T_S(r) = T_{S0}\left[1 - \frac{\epsilon}{4} f(r)\right], \tag{88}$$

where $f(r)$ is given by (71) for deep spots, by (75) for shallow spots, and by (79) for hemispherical spots. Figure 3 shows a plot of the normalized excess temperature in the bright ring,

$$\Delta T_R(r) = \frac{T_S(r) - T_{S0}}{T_S(1) - T_{S0}}, \tag{89}$$

as a function of r, for $D = 0.2$, 1.0, and 5.0. For large r, ΔT_R falls off like $1/r^3$. The width of the region in which ΔT_R is appreciable increases with increasing spot depth. A more precise statement can be made by computing the radius r_w at which $\Delta T_R(r_w) = 0.1$. The results, given in Table III, show that

$$r_w - 1 \approx D \tag{90}$$

for a wide range of values of D. *Thus the width of the bright ring is comparable with the depth of the spot.*

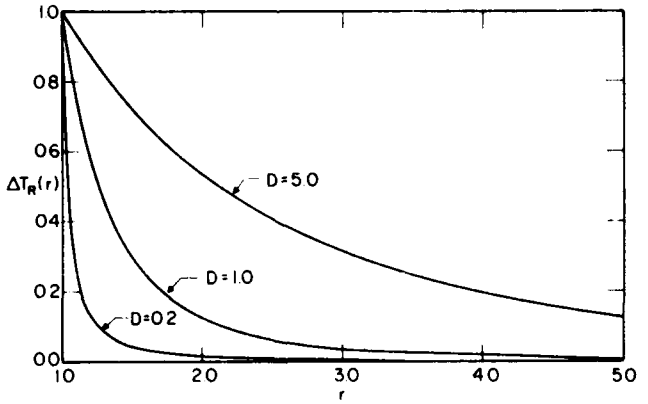


Fig. 3. Temperature excess ΔT_R as a function of radius r in bright ring, normalized to unity at the umbral boundary ($r = 1$). The three curves are for a shallow spot ($D = 0.2$), a hemi-spherical spot ($D = 1.0$), and a deep spot ($D = 5.0$).

TABLE III

The width r_w of the bright ring as a function of spot aspect ratio D

D	$(r_w-1)/D$	D	$(r_w-1)/D$
0.2	1.28	1.5	1.10
0.4	1.26	2.0	1.05
0.6	1.22	3.0	0.99
0.8	1.18	4.0	0.95
1.0	1.15	5.0	0.92

Consider now the amplitude of the bright ring. Here, unfortunately, the linear theory breaks down. One cannot produce an umbral temperature of 4000 K with a small ε. In order to discuss amplitudes, we use the length-scale information obtained above from the linear theory, and we use a simple model which incorporates the correct nonlinear radiative boundary condition on $z = 0$. We do this by assuming a functional form for $T_S(r)$. The peak temperature excess in the bright ring, $\Delta T_p = \beta T_{S0}$, is left as a free parameter, to be determined by the global energy balance. Thus we assume

$$T_S(r) = \begin{cases} T_{U_\infty}, & r < 1 \\ T_{S0}, & r > \left\{1 + \dfrac{\beta}{1 + ((r-1)/w)^3}\right\}. \end{cases} \tag{91}$$

We choose w so that the temperature excess has dropped to $0.1\,\Delta T_p$ for $r = 1 + D$ (in accordance with (90)). This gives

$$w = D/9^{1/3}. \tag{92}$$

We impose the global energy balance:

$$\lim_{R\to\infty} \left\{\sigma\pi T_U^4 + \sigma\pi \int_1^R 2T_S^4 r\,\mathrm{d}r - \sigma\pi T_{S0}^4 R^2\right\} = 0. \tag{93}$$

The substitution of (91) into (93) yields

$$1 - \left(\frac{T_U}{T_{S0}}\right)^4 = 2\beta w(4I_1 + 6\beta I_2 + 4\beta^2 I_3 + \beta^3 I_4) + \\ + 2\beta w^2(4J_1 + 6\beta J_2 + 4\beta^2 J_3 + \beta^3 J_4), \tag{94}$$

where

$$I_n = \int_0^\infty \frac{\mathrm{d}x}{(1+x^3)^n}, \qquad J_n = \int_0^\infty \frac{x\,\mathrm{d}x}{(1+x^3)^n}. \tag{95}$$

By residue theory, one shows easily that

$$(I_1, I_2, I_3, I_4) = \frac{2\pi}{\sqrt{3}}\left(\frac{1}{3}, \frac{2}{9}, \frac{5}{27}, \frac{40}{243}\right), \tag{96}$$

and

$$(J_1, J_2, J_3, J_4) = \frac{2\pi}{\sqrt{3}}\left(\frac{1}{3}, \frac{1}{9}, \frac{2}{27}, \frac{14}{243}\right). \tag{97}$$

From (92), (94), (96), and (97), we obtain a single numerical relation between (T_U/T_{S0}), $\beta = \Delta T_p/T_{S0}$, and D. Figure 4 shows a plot of ΔT_p versus D, for several values of T_U, when $T_{S0} = 5800$ K. The value of ΔT_p is insensitive to the umbral temperature T_U, but very sensitive to the aspect ratio D. It is clear that a shallow spot ($D \leq 1$) in the inhibition theory has an intense bright ring.

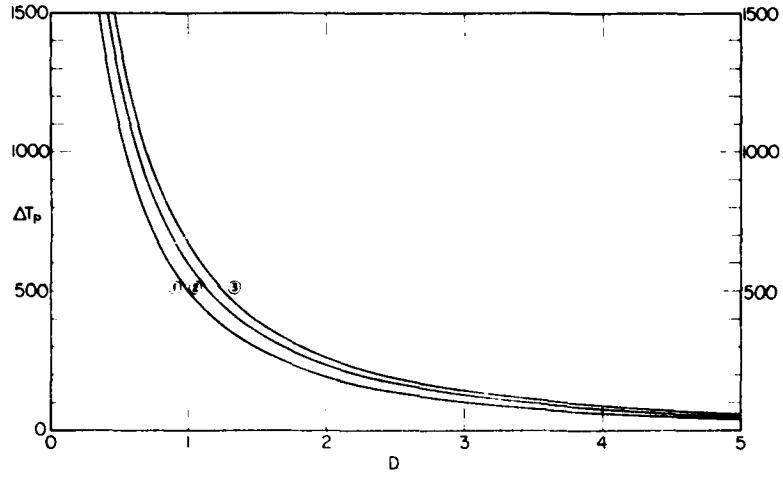


Fig. 4. Peak excess temperature ΔT_p (above photospheric) in the bright ring as a function of spot aspect ratio D, for a photospheric temperature of 5800 K, and an umbral temperature of 4500 K (curve 1), 4000 K (curve 2), and 3500 K (curve 3).

We compare these results with those of other models. Parker (1974a) found intense, narrow bright rings, the width being comparable with the temperature scale height. This is not inconsistent with the present results, since Parker assumed that the spot depth is much less than the temperature scale height. The present work requires the opposite inequality. Although Eschrich and Krause (1977) do not identify the width of the bright ring, nor do they give any numbers for it, their results presented graphically appear to be consistent with the estimate (90) for the width of the bright ring. Their results for the amplitude of the bright ring are somewhat smaller than but quite comparable to our results. Spruit's (1977a) calculations show much weaker

bright rings. He finds, for example, a bright ring with a temperature excess of about 20 K for a cylindrical spot of radius 5000 km and depth 1000 km. As one can show, this implies that the flux missing from the umbra is spread over an area with a diameter of the order of 500 000 km. Spruit attributes this rapid spreading to the strong increase with depth of the conductivity in his model.

Consider now the problem of comparing bright ring models with observation. In principle, the amplitude and scale of the bright ring are observable. In practice, the situation is greatly complicated by the presence of a penumbra in a real spot. For purposes of the present discussion, we will adopt the standard view that the umbra is a measure of the basic, vertical, subsurface flux tube, whereas the penumbra is a vertically thin region in the layers above $\tau = 1$ where the spreading magnetic field is predominantly horizontal. The principal observed fact is that the penumbra is darker than the normal photosphere whereas the models considered here require the greatest temperature excess to be seen just outside the umbra.

There are (at least) two qualitative ways out of this dilemma. First, one can suppose that the excess flux coming up just beneath the penumbra is converted into mechanical energy, which propagates away. Both the Evershed effect and penumbral waves are candidates suggested by observation. This idea cannot be tested observationally until there are detailed predictions about the nature of the mechanical flux, and, most important, about where the mechanical flux is dissipated and becomes visible as excess radiative flux.

A second possibility, considered quantitatively by Spruit (1977a), is that the horizontal magnetic field of the penumbra also inhibits convection. In a thermal model, this would appear as an additional thin region of lowered conductivity. The entire region of lowered conductivity would then have the appearance of an inverted top hat. The difficulty with this picture is that, as Parker (1974a), Eschrich and Krause (1977), and we have shown, thin regions of inhibition produce intense bright rings. In this case, the intense bright ring would appear just outside the penumbra. Whether the rings would be too bright to be consistent with observations cannot be said until detailed model calculations have been carried out. If the effective thermal conductivity of the convection zone increases as rapidly with depth as Spruit (1977a) has suggested, then as his calculations show, the bright ring outside a penumbral model of this sort will indeed be weak.

5.4. The depth of the inhibition region

As the calculations in Section 5.3 show, the thermal inhibition theory, in its present simple form, gives a relation (shown in Figure 4) between umbral temperature, bright ring peak temperature, and spot aspect ratio. Because of the masking effect of the penumbra discussed above, we cannot draw any definite conclusions from the present model about the depth of the inhibition region. However, two of our results suggest that a shallow spot makes greater demands on the theory than a deep spot. First, there is the peak bright ring temperature (Figure 4). The excess flux near the umbral boundary becomes very large as D drops below 1. On the other hand a value

of, say, $D = 5$ poses only a small theoretical problem, since the excess flux corresponds to a maximum excess temperature of the order of 50 K. The second result suggesting deep spots concerns the increased internal spot temperature discussed in Section 5.2. The increase is very large for shallow spots, but is quite small for $D = 5$.

We may ask whether a value $D = 5$ or greater is reasonable from the point of view of dynamics. Weiss (1964) has tabulated, as a function of depth, the magnetic field strength necessary to inhibit convection (the calculation being based on the idea that the magnetic field will seriously interfere with the convection when the magnetic energy density is comparable with the kinetic energy density of the convection). Weiss finds that a field of less than 6000 G will interfere with convection throughout the convection zone. Thus a modest contraction with depth of a flux tube with a surface field of 3000 G is sufficient to give a deep region of inhibition.

5.5. Concluding Remarks

We summarize here the most important features of the present model:

(1) The umbral boundary is sharp even for deep spots. The horizontal scale for temperature variations across the umbral boundary is the superadiabatic temperature scale height at the surface. Since the models of Eschrich and Krause (1977) and Spruit (1977a) also show a sharp umbral boundary, it seems safe to conclude that it is a general consequence of the inhibition theory.

(2) The missing umbral flux shows up in the model as a bright ring around the umbra. Since this is a consequence of flux conservation, it is a general feature of the inhibition theory. The intensity of the bright ring, however, is highly model dependent. In our model, in the model of Eschrich and Krause (1977), and in Parker's model (1974a), shallow spots give intense bright rings. Spruit (1977a), however, finds weak bright rings even for shallow spots. He attributes this to the strong horizontal spreading of flux disturbances associated with the depth-dependent conductivity in his models. All of the models agree in predicting weak bright rings for sufficiently deep spots.

(3) Only a thin surface layer of the spot is cool. At a depth of several superadiabatic scale heights or more, the material directly beneath the visible umbra is hotter than the surrounding normal atmosphere at the same depth. Although Eschrich and Krause (1977) do not discuss this point, Spruit (private communication) finds similar results, and it is probable that the subsurface heating is a general feature of the inhibition theory.

In spite of the different descriptions of turbulent transport used by Eschrich and Krause (1977), by Spruit (1977a), and in the present work, all of the models yield surface temperature distributions resembling sunspots. This provides support for the inhibition theory as an explanation of why sunspots are dark.

Acknowledgements

I wish to thank Professors P. A. Clark and J. H. Thomas for many helpful discussions

of this work, and for reading the final manuscript. I thank Dr H. C. Spruit for his useful commnets on the final manuscript. I thank Professors R. F. Gans, H. L. Helfer, J. E. Molyneux, M. P. Savedoff, and H. Van Horn for helpful comments on earlier presentations of this work. Mr Mark Scheuer helped with some of the computational problems.

This work was supported in part by Air Force Contract F19628-77-C-0079 through Sacramento Peak Observatory.

References

Abramowitz, M. and Stegun, I. A.: 1964, *Handbook of Mathematical Functions*, U.S. Government Printing office, Washington.
Biermann, L.: 1941, *Vierteljahrsschr. Astron. Gesellsch.* **76**, 194.
Boruta, N.: 1977, *Astrophys. J.* **215**, 364.
Bray, R. J. and Loughhead, R. E.: 1965, *Sunspots*, John Wiley, New York.
Busse, F. H.: 1973, *Solar Phys.* **33**, 413.
Chitre, S.: 1963, *Monthly Notices Roy. Astron. Soc.* **126**, 431.
Chitre, S. and Shaviv, G.: 1967, *Solar Phys.* **2**, 150.
Cowling, T. G.: 1953, in G. P. Kuiper (ed.), *The Sun*, University of Chicago Press, Chicago.
Cowling, T. G.: 1976a, *Monthly Notices Roy. Astron. Soc.* **177**, 409.
Cowling, T. G.: 1976b, *Magnetohydrodynamics*, Adam Hilger, Bristol.
Danielson, R. E.: 1965, in R. Lüst (ed.), 'Stellar and Solar Magnetic Fields', *IAU Symp.* **22**, 314.
Deinzer, W.: 1965, *Astrophys. J.* **141**, 548.
Eschrich, K.-O. and Krause, F.: 1977, *Astron. Nachr.* **298**, 1.
Galloway, D. J., Proctor, M. R. E., and Weiss, N. O.: 1977, *Nature* **266**, 686.
Hastings, C.: 1955, *Approximations for Digital Computers*, Princeton University Press, Princeton.
Hoyle, F.: 1949, *Some Recent Researches in Solar Physics*, Cambridge University Press, Cambridge.
Meyer, F., Schmidt, H. U., Weiss, N. O., and Wilson, P. R.: 1974, *Monthly Notices Roy. Astron. Soc.* **169**, 35.
Morse, P. M. and Feshbach, H.: 1953, *Methods of Theoretical Physics II*, McGraw-Hill, New York.
Mullan, D. J.: 1973, *Solar Phys.* **30**, 75.
Mullan, D. J.: 1974, *Astrophys. J.* **187**, 621.
Parker, E. N.: 1955, *Astrophys. J.* **121**, 491.
Parker, E. N.: 1974a, *Solar Phys.* **36**, 249.
Parker, E. N.: 1974b, *Solar Phys.* **37**, 127.
Parker, E. N.: 1975a, *Solar Phys.* **40**, 275.
Parker, E. N.: 1975b, *Solar Phys.* **40**, 291.
Parker, E. N.: 1976, *Astrophys. J.* **204**, 259.
Parker, E. N.: 1977, *Monthly Notices Roy. Astron. Soc.* **179**, 93P.
Roberts, B.: 1976, *Astrophys. J.* **204**, 268.
Spruit, H. C.: 1977a, *Solar Phys.* **55**, 3.
Spruit, H. C.: 1977b, Ph.D. Thesis, Utrecht.
Thomas, J. H.: 1978, *Astrophys. J.* **225**, 275.
Van Dyke, M.: 1975, *Perturbation Methods in Fluid Mechanics*, Parabolic Press, Stanford.
Weiss, N. O.: 1964, *Monthly Notices Roy. Astron. Soc.* **128**, 225.
Weiss, N. O.: 1969, in D. G. Wentzel and D. A. Tidman (eds.), *Plasma Instabilities in Astrophysics*, Gordon and Breach, New York.
Yun, H. S.: 1970, *Astrophys. J.* **162**, 975.

Variations of the Sun's radius and temperature due to magnetic buoyancy

LIVINGSTON[1] has recently measured a decrease in the surface temperature of the Sun coincident with increased solar activity. He interpreted the temperature drop as implying a corresponding reduction in luminosity. I point out here that surface cooling could also be due to a radial expansion of the Sun, with no attendant reduction in luminosity. There is a plausible physical mechanism for such an expansion; namely, variations in magnetic buoyancy due to variations in the magnetic flux in the convection zone over the solar cycle.

The solar luminosity L may be expressed in terms of an effective surface temperature T as $L = 4\pi R^2 \sigma T^4$, where R is the solar radius. For constant R, a decrease in T implies a decrease in L. Alternatively, there can be a cooling (or heating) of the surface due to a radial expansion (or contraction) of the Sun with constant L. For small changes ΔT and ΔR with L constant,

$$\frac{\Delta R}{R} = -2\frac{\Delta T}{T} \tag{1}$$

As $T \sim 6{,}000$ K, a drop in surface temperature of 1 K would correspond to a relative expansion $\Delta R/R \sim 3 \times 10^{-4}$.

The concept of magnetic buoyancy was introduced by Parker[2] and Jensen[3] and is an important ingredient in the solar dynamo. Consider an isolated magnetic flux tube of field strength B in the solar interior. The sum of the gas pressure p_i and the magnetic pressure $p_m = B^2/8\pi$ inside the tube must balance the external gas pressure p_e, so $p_i + p_m = p_e$, and $p_i < p_e$. If the flux tube is in thermal equilibrium with its surroundings ($T_i = T_e$), then the density inside the tube is lower than the density of the surroundings ($\rho_i < \rho_e$) and the flux tube is buoyant. Jensen[3] pointed out that, because of the density depression inside a magnetic flux tube, the presence of magnetic flux tubes in the solar interior will cause the volume of the Sun to increase; thus, we would expect the radius of the Sun to vary with the solar cycle, with maximum radius occurring near sunspot maximum. Jensen's suggestion, which has attracted little previous attention, is pursued here.

How much expansion and contraction of the Sun might we expect due to variations in magnetic buoyancy over the solar cycle? A precise estimate would require detailed knowledge of the magnetic field distribution in the convection zone over the solar cycle and a complete calculation of the effect of the magnetic field on the structure of the solar envelope. This is beyond present capabilities, so instead a crude calculation is used to get a very rough estimate. Consider a toroidal flux tube that circles the Sun at spherical radius R_1 and constant latitude ϕ. Let the tube have field strength B and cross-sectional radius a. The mass of gas within the tube is less than would occupy the same volume in the absence of the magnetic field. The difference in densities outside and inside the tube is given by

$$\Delta\rho = \rho_e - \rho_i = (p_e - p_i)/RT_i = p_m/RT_i$$

where $T_i = T_e = T_i$ is the temperature of the flux tube and its immediate surroundings. The total mass defect in the tube is equal to the volume of the tube times $\Delta\rho$, or

$$\Delta M = 2\pi R_1 \cos\phi\, \pi a^2 p_m/RT_i$$

Assuming that this mass ΔM is spread over a thin spherical shell of mean radius R_1 and thickness ΔR, then the mass in this shell is given by

$$\Delta M = 4\pi R_1^2 \Delta R \rho_e$$

Equating the two expressions for ΔM gives

$$\frac{\Delta R}{R} = \frac{\pi}{2}\cos\phi \left(\frac{R}{R_1}\right)\left(\frac{a}{R}\right)^2 \frac{p_m}{p_e} \tag{2}$$

This expression gives an estimate of the relative change in radius due to a single toroidal flux tube.

To obtain a numerical estimate we can, for example, use the results of the study by Weiss[4] of magnetic flux tubes in the solar convection zone, based on equipartition of magnetic energy and kinetic energy of convection. Weiss finds that a typical flux tube has a total flux of 4×10^{21} Mx. The radius and field strength of the flux tube vary slowly from 5,000 km and 5,700 G at the bottom of the convection zone to 14,000 km and 600 G at photosphere. The ratio p_m/p_e varies greatly, from 6×10^{-6} at the bottom of the convection zone to 2.8×10^{-1} at the photosphere. Thus, the dominant contribution to ΔR will come from flux tubes near the top of the convection zone. Using values given by Weiss for a and p_m/p_e in equation (2), for R_1/R in the range 0.98–1.0, we obtain values of $\Delta R/R$ in the range $5 \times 10^{-8} \cos\phi$ to $2 \times 10^{-4} \cos\phi$. For example, a single flux tube at a depth of 1,000 km at middle latitude gives $\Delta R/R \sim 2 \times 10^{-5}$. Now, to estimate the change in solar radius with the solar cycle, we need to estimate the difference in total magnetic flux in the upper convection zone between solar maximum and minimum. If we assume that the total magnetic flux emerging at the solar surface is a good indicator of the total flux in the uppermost layers of the convection zone, then observations suggest a difference in total flux of the order 10^{23} Mx between solar maximum and minimum. This is equivalent to some 25 or more of the individual flux tubes discussed above and could imply a relative change in radius $\Delta R/R \sim 5 \times 10^{-4}$ or more and a corresponding drop in surface temperature of ~ 1.5 K or more.

Magnetic field strengths in the convection zone may be higher than the equipartition limit used in the estimate above. Indeed, observations[5] have shown that almost all of the magnetic flux in the quiet photosphere is concentrated to strengths of 1,000–2,000 G, well above the equipartition limit, and the theory of Galloway, Proctor and Weiss[6] predicts that magnetic fields will be concentrated to 10 times the equipartition limit or more by the convection. Equation (2) can be rewritten as

$$\frac{\Delta R}{R} \approx \frac{FB\cos\phi}{16\pi R R_1 p_e}$$

where $F = \pi a^2 B$ is the total magnetic flux in the tube. For a tube of fixed total flux, the relative change in radius is proportional to the magnetic field strength. Thus, higher field strengths in the convection zone would increase the estimate of $\Delta R/R$ for a fixed difference in total flux over the solar cycle. With higher field strengths, flux tubes deeper in the convection zone would contribute significantly to $\Delta R/R$. Although there is uncertainty about magnetic fields in the convection zone, the estimates do suggest that the surface cooling of ~ 6 K observed by Livingston may be at least partly caused by expansion due to magnetic buoyancy.

The historical record of measurements of the solar radius gives no clear or consistent picture[7-9]. The currently accepted value[10] ($R = 959.63''$ at 1 AU) dates back to 1891 and is subject to correction for irradiation[11]. Variability of R with amplitudes ranging from $\sim 0.05''$ to $\sim 2''$ and periods of 7, 8, 11 and 22 yr has been reported. Giannuzzi[12], for example, found variations in R with amplitude 0.2–0.5″ and period 22 yr. Meyermann[13] reported variations with total amplitude $\sim 0.15''$ ($\Delta R/R \sim 1.5 \times 10^{-4}$) and period 11 yr with maximum radius occurring very near sunspot maximum; this is in good agreement with the theoretical argument above. Gething[8] finds variations in R with amplitude ~ 0.1–$0.2''$, but with no obvious correlation with the solar cycle. Improved measurements of the solar radius over a solar cycle are needed to test the ideas expressed here. Recent improvements[9,14] in the theoretical definition of the solar radius should give a considerable increase in accuracy. Changes in opacity during expansion may affect the visual change in solar radius. Note that the solar oblateness measured by Dicke and Goldenberg[15] ($\Delta R/R \sim 5 \times 10^{-5}$) and later disputed by Hill and Stebbins[16] is an order of magnitude smaller than the variation in mean radius estimated above. Even if the expansion due to magnetic buoyancy is too small to produce a measurable change in surface temperature, it may still be large enough to show up directly as a change in radius. Accurate monitoring of the solar

radius over the solar cycle would give a measure of the variation of total magnetic flux in the convection zone and add to our understanding of the solar dynamo. Measurements of the solar radius from space are particularly desirable.

This research was supported by grants from the US Air Force Geophysics Laboratory and NASA. I thank Alfred Clark, Jr and Eugene Parker for helpful comments.

JOHN H. THOMAS

Department of Mechanical and Aerospace Sciences
and
CEK Mees Observatory,
University of Rochester,
Rochester, New York 14627

Received 6 April; accepted 10 July 1979.

1. Livingston, W. C. *Nature* **272**, 340–341 (1978).
2. Parker, E. N. *Astrophys. J.* **121**, 491–507 (1955).
3. Jensen, E. *Ann. Astrophys.* **18**, 127–140 (1955).
4. Weiss, N. O. *Mon. Not. R. astr. Soc.* **128**, 225–235 (1964).
5. Harvey, J. W. in *Highlights of Astronomy* (ed. Müller, E.) Vol. 4, 223–239 (Reidel, Dordrecht, 1977).
6. Galloway, D. J., Proctor, M. R. E. & Weiss, N. O. *Nature* **266**, 686–689 (1977).
7. Goldberg, L. in *The Sun* (ed. Kuiper, G. P.) 1–35 (University of Chicago Press, 1953).
8. Gething, P. J. D. *Mon. Not. R. astr. Soc.* **115**, 558–570 (1955).
9. Wittmann, A. *Solar Phys.* **29**, 333–340 (1973).
10. Auwers, A. *Astr. Nachr.* **128**, 367 (1891).
11. Cullen, R. T. *Mon. Not. R. astr. Soc.* **86**, 344–349 (1926).
12. Giannuzzi, M. A. *Contr. Sci. Oss. Roma*, No. 211 (1955).
13. Meyermann, B. *Astr. Nachr.* **279**, 45–46 (1950).
14. Hill, H. A., Stebbins, R. T. & Oleson, J. R. *Astrophys. J.* **200**, 484–498 (1975).
15. Dicke, R. H. & Goldenberg, H. M. *Phys. Rev. Lett.* **18**, 313–316 (1967).
16. Hill, H. A. & Stebbins, R. T. *Astrophys. J.* **200**, 471–483 (1975).

(Reprinted from Nature Vol. 280, No. 5724, pp. 662-663, August 23 1979)

Printed in Great Britain by Macmillan Production Ltd., Basingstoke, Hampshire

UMBRAL OSCILLATIONS AS RESONANT MODES OF MAGNETO-ATMOSPHERIC WAVES

Mark A. Scheuer and John H. Thomas

Department of Mechanical and Aerospace Sciences
and
C.E. Kenneth Mees Observatory
University of Rochester

Abstract

Umbral oscillations in sunspots are identified as a resonant response of the umbral atmosphere to forcing by oscillatory convection in the subphotosphere. The full, linearized equations for magneto-atmospheric waves are solved numerically for a detailed model of the umbral atmosphere, for both forced and free oscillations. Resonant "fast" modes are found, the lowest mode having a period of 153 s, typical of umbral oscillations. A comparison is made with a similar analysis by Uchida and Sakurai (1975), who calculated resonant modes using an approximate ("quasi-Alfvén") form of the wave equations. Whereas both analyses give an appropriate value for the period of oscillation, several new features of the motion follow from the full equations.

1. Introduction

Observations of umbral oscillations in sunspots (Beckers and Schultz 1972; Bhatnagar and Tanaka 1972; Giovanelli 1972; Rice and Gaizauskas 1973; Phillis 1975; Moore and Tang 1975) give a fairly consistent picture of the oscillations as a resonant wave mode in the umbra. Doppler shifts in Hα and in photospheric lines show a continuous periodic oscillation in sunspot umbrae, with a fairly well defined period in the range 145-190 s.

Moore (1973) has argued that the driving mechanism for umbral oscillations is overstable oscillatory convection in a shallow subphotospheric layer in the umbra (see also Mullan and Yun 1973). Uchida and Sakurai (1975) also consider the driving mechanism to be overstable convection, and they interpret the umbral oscillations as a resonant response of the umbral atmosphere to this forcing. Their resonant mode is a standing quasi-Alfvén wave trapped in the photosphere and chromosphere. In a similar approach, Antia and Chitre (1979) identify the umbral oscillations with an overstable fast magneto-atmospheric wave mode in the umbra. Their calculations account more fully for the effects of compressibility.

In this paper, we adopt a point of view very close to that of Uchida and Sakurai (1975). We shall argue that the umbral oscillations are a resonant response to overstable convection in a thin subphotospheric layer and that the resonant wave mode

is a fast magneto-atmospheric wave. Our basic model atmosphere is similar to Uchida and Sakurai's. There are, however, several fundamental differences between our work and Uchida and Sakurai's which lead to a new understanding of the umbral oscillations. These differences include the following:

(1) We do not adopt the approximation that leads to the quasi-Alfvén wave; rather, we solve the complete linearized magneto-atmospheric wave equations. We show that, although the quasi-Alfvén approximation yields a good estimate of the period of oscillation, it fails to describe certain important features of the motion.

(2) Instead of assuming a reflecting lower boundary at the base of the photosphere, we include a semi-infinite lower layer in our model atmosphere to represent the umbral convection zone. We show that, as far as umbral oscillations are concerned, this lower layer acts much like a reflecting boundary, but not for the reasons given by Uchida and Sakurai.

(3) In addition to calculating free eigenmodes of oscillation in our model atmosphere, we also calculate the response to forcing at different frequencies.

(4) We show that the downward reflection of wave energy is not total, but that a small fraction of energy escapes into the corona in the form of acoustic waves along the magnetic field lines.

2. Basic Equations and the Umbral Model

In our simplified treatment of a sunspot umbra, we consider the undisturbed magnetic field to be uniform and vertical, $\underset{\sim}{B} = (0,0,B_o)$. The atmosphere is assumed to be a compressible, inviscid, perfectly conducting gas under uniform gravity $g(= 0.274 \text{ km s}^{-2})$ in the negative z-direction. The undisturbed pressure $p(z)$, density $\rho(z)$, and temperature $T(z)$ are functions of height z only, and hydrostatic equilibrium requires that

$$\frac{dp}{dz} = -\rho g \ .$$

We then consider small adiabatic perturbations of this equilibrium atmosphere. There is no preferred horizontal direction, so in cartesian coordinates we may take the horizontal wavenumber in the x-direction and assume that the perturbation velocity $\underset{\sim}{u} = (u,v,w)$ has the form $\underset{\sim}{u} = \hat{\underset{\sim}{u}}(z)\exp[i(kx - \omega t)]$, with $\hat{\underset{\sim}{u}}(z) = [\hat{u}(z), \hat{v}(z), \hat{w}(z)]$. Starting with the linearized equations of induction and conservation of mass, momentum, and energy, we can eliminate the perturbations in pressure, density, and magnetic field and arrive at a single vector equation for the perturbation velocity (cf. Ferraro and Plumpton 1958). The y component of this equation is

$$[v_A^2 \frac{d^2}{dz^2} + \omega^2]\hat{v} = 0 \ , \qquad (1)$$

where $v_A = (B_o^2/4\pi\rho(z))^{\frac{1}{2}}$ is the Alfvén speed. This equation is decoupled from the other two components and represents a pure, transverse Alfvén wave with horizontal, incompressible motion. The x and z component equations are

$$[v_A^2(\frac{d^2}{dz^2} - k^2) - c^2k^2 + \omega^2]\hat{u} + ik[c^2 \frac{d}{dz} - g]\hat{w} = 0, \tag{2}$$

$$[c^2 \frac{d^2}{dz^2} - \gamma g \frac{d}{dz} + \omega^2]\hat{w} + ik[c^2 \frac{d}{dz} - (\gamma-1)g]\hat{u} = 0, \tag{3}$$

where $c = (\gamma RT(z))^{\frac{1}{2}}$ is the sound speed and γ is the ratio of specific heats. Equations (2) and (3) represent fully coupled magneto-atmospheric waves in which compression, buoyancy, and magnetic forces all play a role.

Equations (2) and (3) also describe the vertical dependence of waves in cylindrical coordinates (r,θ,z). If we take the perturbation velocity $\underset{\sim}{u} = (u_r,u_\theta,u_z)$ in the form

$$u_r = \hat{u}(z)[kJ_{m+1}(kr) - \frac{m}{r} J_m(kr)]\exp[i(m\theta - \omega t)], \tag{4}$$

$$u_\theta = -im\ \hat{u}(z)\ \frac{J_m(kr)}{r} \exp[i(m\theta - \omega t)], \tag{5}$$

$$u_z = -ik\ \hat{w}(z)J_m(kr)\exp[i(m\theta - \omega t)], \tag{6}$$

we arrive at equations identical to (2) and (3), except that in this case k is a radial wavenumber. We shall consider only axisymmetric modes $(m=0)$, for which $u_\theta = 0$ and

$$u_r = k\hat{u}(z)J_1(kr)\exp(-i\omega t), \tag{7}$$

$$u_z = -ik\,\hat{w}(z)J_o(kr)\exp(-i\omega t). \tag{8}$$

(There are also axisymmetric modes with $u_\theta \neq 0$, representing pure, torsional Alfvén waves; these are not represented by the form (4)-(6).) If we require, as Uchida and Sakurai do, that the radial velocity u_r vanish at the edge of the umbra, $r = a$, then there is a discrete set of radial wavenumbers k_j defined by $J_1(k_j a) = 0$. We shall consider only the first zero of J_1, $ka = 3.83$, and take the same value $k^{-1} = 1100$ km as Uchida and Sakurai, corresponding to an umbral radius $a = 4200$ km.

The temperature distribution $T(z)$, and hence the density distribution $\rho(z)$, must be specified to represent the umbral atmosphere. We adopt as our model umbra the three-layer atmosphere shown in Figure 1. This same atmosphere was used in a previous study of reflection of Alfvén waves in the umbra (Thomas 1978), and the upper two layers are identical to the model atmosphere of Uchida and Sakurai. Layer 2, representing the umbral photosphere and chromosphere, is isothermal at temperature T_o. Layer 3 represents the corona and is isothermal at temperature T_1 ($>T_o$). The transition region is represented as a discontinuity in temperature (and density) at height $z = z_t$. Layer 1, which represents the umbral convection zone, is assumed to have a linear temperature distribution $T(z) = T_o - \beta z$

with $\beta > 0$. The value of the temperature gradient β may be chosen equal to or slightly greater than the adiabatic gradient $\beta_s = g/c_p$. Let $\alpha = H_o\beta/T_o$ be a nondimensional measure of the temperature gradient β in layer 1, where $H_o = RT_o/g$ is the density scale height in layer 2. The temperature, density, sound speed, and Alfvén speed in each layer are then given as follows.

Layer 1 $(z < 0)$:

$$\begin{aligned} T(z) &= T_o\left(1 - \frac{\alpha z}{H_o}\right), \quad H_o = RT_o/g, \\ \rho(z) &= \rho_o\left(1 - \frac{\alpha z}{H_o}\right)^{\frac{1-\alpha}{\alpha}}, \\ c^2(z) &= c_o^2\left(1 - \frac{\alpha z}{H_o}\right), \quad c_o^2 = \gamma RT_o, \\ v_A^2(z) &= v_{Ao}^2\left(1 - \frac{\alpha z}{H_o}\right)^{\frac{\alpha-1}{\alpha}}, \quad v_{Ao}^2 = B_o^2/4\pi\rho_o. \end{aligned} \tag{9}$$

Layer 2 $(0 \le z \le z_t)$:

$$\begin{aligned} T(z) &= T_o, \quad H_o = RT_o/g, \\ \rho(z) &= \rho_o \exp(-z/H_o), \\ c^2(z) &= c_o^2 = \gamma RT_o, \\ v_A^2(z) &= v_{Ao}^2 \exp(z/H_o), \quad v_{Ao}^2 = B_o^2/4\pi\rho_o. \end{aligned} \tag{10}$$

Layer 3 $(z > z_t)$:

$$
\begin{aligned}
T(z) &= T_1, \quad H_1 = RT_1/g \\
\rho(z) &= \rho_o \left(\frac{T_o}{T_1}\right) \exp[-z_t/H_o - (z-z_t)/H_1], \\
c^2(z) &= c_1^2 = \gamma R T_1, \\
v_A^2(z) &= v_{Ao}^2 \left(\frac{T_1}{T_o}\right) \exp[z_t/H_o + (z-z_t)/H_1].
\end{aligned}
\tag{11}
$$

We note here that, although we do compute solutions of the wave equations (2) and (3) in the entire three-layer atmosphere, we also present results for simpler versions of the atmosphere, where the lower layer is replaced by a reflecting boundary or where the upper layer is omitted and layer 2 extends to $z = \infty$.

For comparison with observations, we adopt the following numerical values for the parameters in our umbral model: $T_o = 4500$ K, $T_1 = 2\times10^6$ K, $B_o = 1000$ G, $\rho_o = 5\times10^{-7}$ g cm^{-3}, $\gamma = 5/3$, and $\alpha = 0.5$. The sound speed and density scale height in layer 2 are then $c_o = 7.9$ km s^{-1} and $H_o = 136.5$ km, and the Alfvén speed at $z = 0$ is $v_{Ao} = 4.0$ km s^{-1}. We take the height of the transition region to be $z_t = 20\ H_o = 2730$ km.

3. Analysis

In order to compute wave modes in our model umbral atmosphere, we need to solve the basic coupled wave equations (2)

and (3) for two specific cases, uniform temperature (layers 2 and 3) and linearly varying temperature (layer 1). Consider first the isothermal case. The problem of magneto-atmospheric waves in an isothermal atmosphere with a uniform vertical magnetic field was first treated by Ferraro and Plumpton (1958) in an important early paper that has escaped notice in subsequent treatments of this problem by Uchida and Sakurai (1975) and Hollweg (1979). Ferraro and Plumpton solve the full wave equations (2) and (3), whereas Uchida and Sakurai and Hollweg solve approximate forms of the equations which hold only in the case $c^2 << v_A^2$. This approximation is not good in the low photosphere, so we prefer to use the full equations. In contrast to the case of a uniform horizontal magnetic field (Nye and Thomas 1976), the wave equations for an isothermal atmosphere with a uniform vertical magnetic field do not have solutions expressible in terms of tabled special functions. Ferraro and Plumpton construct power series solutions to the equations and give only limited numerical results. For our purposes, it is more convenient to rely on direct numerical integration of the equations.

It is useful to have asymptotic solutions of the wave equations for large z. Suppose the transition zone is absent and layer 2 extends to $z = \infty$. Then, for large z, $v_A^2 >> c^2$, and the asymptotic solutions of Equations (2) and (3) for $z \to \infty$ are (cf. Ferraro and Plumpton 1958)

$$\hat{u}(z) \sim A \exp(-Kz/H_o) + B \exp(-z/2H_o)\exp[\pm i(\Omega^2-\tfrac{1}{4})^{\frac{1}{2}} z/H_o], \quad (12)$$

$$\hat{w}(z) \sim C \exp(z/2H_o)\exp[\pm i(\Omega^2-\tfrac{1}{4})^{\frac{1}{2}} z/H_o], \quad (13)$$

where $K = H_o k$ is a nondimensional horizontal wavenumber and $\Omega = H_o\omega/c_o$ is a nondimensional frequency. For $\Omega > \frac{1}{2}$ the plus sign in the exponents corresponds to an upward propagating wave and the negative sign to a downward propagating wave. The acoustic cutoff frequency corresponds to $\Omega = \frac{1}{2}$, and for $\Omega < \frac{1}{2}$ the asymptotic solution represents purely evanescent behavior.

The case of a linear temperature variation (layer 1) has not been studied previously. We have not found a useful representation of the solution to the wave equations in terms of special functions, so we again use direct numerical integration.

In analyzing our computed solutions, it is useful to consider the distribution of wave energy density with height. The total energy density E is given by

$$E = \tfrac{1}{2}\rho\hat{u}^2 + \tfrac{1}{2}\rho\hat{w}^2 + \tfrac{1}{2}\frac{\rho}{\Omega^2}[K^2\hat{u}^2 + (\hat{w}-H_o\frac{d\hat{w}}{dz})^2 + (\frac{\gamma-1}{\gamma})^2\hat{w}^2] + \tfrac{1}{2}\rho(\frac{H_o N^2}{\gamma g\Omega^2})\hat{w}^2 + \tfrac{1}{2}\rho\frac{v_A^2}{c^2}[H_o^2(\frac{d\hat{u}}{dz})^2 + K^2\hat{u}^2] \quad (14)$$

where

$$N^2 = (\frac{\gamma-1}{\gamma})\frac{g}{H_o} = (\gamma-1)\frac{g^2}{c_o^2}$$

in layer 2,

$$N^2 = \left(\frac{\gamma-1}{\gamma}\right)\frac{g}{H_1} = (\gamma-1)\,\frac{g^2}{c_1^2}$$

in layer 3, and

$$N^2 = \left[\frac{\gamma-\alpha\gamma-1}{\gamma\left(1-\frac{\alpha z}{H_o}\right)}\right]\frac{g}{H_o}$$

in layer 1. The five terms on the right hand side of equation (14) represent, in order, the kinetic energy density due to horizontal motions, the kinetic energy density due to vertical motions, the potential energy density associated with adiabatic compression or expansion, the potential energy density associated with the buoyant force, and the potential energy density stored in the magnetic field perturbation.

The wave equations (2) and (3) are solved numerically using a generalized Newton-Raphson relaxation method. The equations are written as a set of linear, first-order ordinary differential equations with the velocities (and eigenvalues, if desired) expressed as some trial value plus a correction. They are then linearized with respect to the corrections. The atmosphere is divided into a grid of equally spaced points. Grid spacings between $0.1\ H_o$ and $0.25\ H_o$ give sufficient accuracy in all of the calculations. Using a finite difference scheme, the calculation is started at one boundary, and, using the equations, the corrections to the initial trial values are calculated at each internal grid point. The conditions at the

other boundary close the set of equations for the necessary corrections at every mesh point. The corrections are then added to the velocities (and eigenvalues) giving the trial quantities for the next iteration. The process is repeated until the maximum correction divided by the associated quantity is less than some pre-assigned value (0.1%).

4. Forced Oscillations

To simulate the forcing by overstable convection in a thin subphotospheric layer we apply an oscillation of fixed frequency and horizontal wavenumber and of unit amplitude on the plane $z = 0$. We consider separately the forcing due to a purely horizontal motion $(\hat{u}(0) = 1, \hat{w}(0) = 0)$ and a purely vertical motion $(\hat{u}(0) = 0, \hat{w}(0) = 1)$ at $z = 0$. A true representation of the forced oscillations will then be some linear combination of these two forced solutions.

We shall consider these forced solutions in two stages. First we consider only a two-layer atmosphere consisting of layers 1 and 2, with layer 2 extending to $z = \infty$ and with layer 3 (the corona) absent (i.e., $z_t \to \infty$). Then we consider solutions for the full three-layer atmosphere. This approach will elucidate the role of the chromosphere - corona transition in providing downward reflection of wave energy.

4.1 TWO-LAYER MODEL

Here layer 2 is semi-infinite $(z_t \to \infty)$ and layer 3 (the

corona) is absent. In layer 2 we assume that the asymptotic solution (12),(13) is valid above $z = 20\ H_o$ and use this as an end point of our numerical integration. At $z = 20\ H_o$ we apply an outward radiation condition, that is, we allow only outward propagation (in the positive z-direction) of energy. This leads to the following matching conditions at $z = 20\ H_o$:

$$\frac{d\hat{u}}{dz} - \frac{1}{H_o}\left[\frac{\exp(-Kz/H_o) + [i(\Omega^2-\tfrac{1}{4})^{\frac{1}{2}} - \tfrac{1}{2}]\exp(-z/2H_o)\exp[i(\Omega^2-\tfrac{1}{4})^{\frac{1}{2}}z/H_o]}{\exp(-z/2H_o)\exp[i(\Omega^2-\tfrac{1}{4})^{\frac{1}{2}}z/H_o] - \frac{1}{K}\exp(-Kz/H_o)}\right]\hat{u} = 0,$$

$$\frac{d\hat{w}}{dz} - \frac{1}{H_o}[i(\Omega^2-\tfrac{1}{4})^{\frac{1}{2}} + \tfrac{1}{2}]\hat{w} = 0. \qquad (15)$$

The solution in layer 1 is more difficult to obtain because there is no simple asymptotic form in the region of interest. We obtain a simple condition of outward propagation (in the negative z-direction) by considering the kinetic energy density, which must remain finite as $z \to -\infty$. Since the density is increasing linearly (for $\alpha = \frac{1}{2}$), the velocity components must drop off at least as fast as $z^{-\frac{1}{2}}$. Consequently, the derivatives must drop off even faster so we apply the conditions

$$\frac{d\hat{u}}{dz} = 0\ , \qquad \frac{d\hat{w}}{dz} = 0 \qquad (16)$$

at some depth $z = -h$. To insure that the error introduced by this is smaller than our allowed numerical error (0.1%), the value of h is increased until changes in h have no effect

on $\hat{u}$ and $\hat{w}$. We found such a suitable value to be $h = -20\ H_o$.

In summary, we evaluate numerically the solution for either horizontal $(\hat{u}(0) = 1,\ \hat{w}(0) = 0)$ or vertical $(\hat{u}(0) = 0,\ \hat{w}(0) = 1)$ forcing, subject to the outward radiation conditions (15) and (16), applied at $z = 20\ H_o$ and $z = -20\ H_o$, respectively. In the resulting solutions, $d\hat{u}/dz$ and $d\hat{w}/dz$, and hence the pressure and magnetic field perturbations, are discontinuous across $z = 0$. This is acceptable in view of the fact that the plane $z = 0$ represents an overstable layer of finite thickness.

In the case of horizontal forcing $(\hat{u}(0) = 1,\ \hat{w}(0) = 0)$, we have computed solutions for a fixed wavenumber $K = 0.124$ (corresponding to $k^{-1} = 1100$ km and $H_o = 136.5$ km) and for many different frequencies. Figure 2 shows a plot of the peak energy density (which occurs at some $z \geq 0$) of the forced oscillation as a function of the forcing frequency Ω. This graph clearly shows the existence of a resonant response at certain frequencies. Numerical values of the first three resonant frequencies Ω_R determined in Figure 2 are listed in Table 1, along with the corresponding dimensional periods using the numerical values of atmospheric parameters listed in Section 2. Each of these frequences lies above the acoustic cutoff frequency, $\Omega_c = 1/2$ $(\omega_c = c_o/2H_o)$ and the resonant modes are "fast" magneto-atmospheric waves. Similar calculations for the case of vertical forcing show a resonant response

at the same resonant frequencies as in the case of horizontal forcing. The period of the lowest resonance, 153 s, lies comfortably in the range of observed periods of umbral oscillations.

Figure 3 shows the solution for the horizontal velocity $\hat{u}(z)$ as a function of height in layer 2 for each of the first three resonant frequencies. The computed magnitude is finite in each case because the resonant frequency is only approximately matched. These first three resonant modes closely resemble the corresponding free eigenmodes computed by Uchida and Sakurai (1975, Figure 3) for their model; we shall say more about this correspondence in Section 5. Each higher resonance has an additional node in the solution for $\hat{u}(z)$. The numerical solutions for $\hat{u}(z)$ agree with the asymptotic solution to within 0.1% for $z \gtrsim 6H_o$; this more than justifies our matching to the asymptotic solution at $z = 20H_o$.

The horizontal velocity $\hat{u}(z)$ tends toward zero as $z \to \infty$, due primarily to the exponentially increasing Alfvén speed. The vertical velocity, however, grows with height, eventually increasing exponentially according to its asymptotic form (13). Because of this rapid growth it is more convenient to plot the logarithmic derivative $H_o(d\hat{w}/dz)/\hat{w}$ than $\hat{w}$ in this layer. Figure 4 shows the computed magnitude of this logarithmic derivative as a function of height is the upper layer, for horizontal forcing at the

lowest resonant frequency. We see that $|H_o(d\hat{w}/dz)/\hat{w}|$ attains its asymptotic value Ω (cf. Equation (13)) for $z \gtrsim 10H_o$.

The behavior of the vertical velocity in this case of purely horizontal forcing is of some interest. Although vertical motions are not forced directly ($\hat{w}(0) = 0$), they are produced above and below $z = 0$ by the "squeezing" effect of the divergent horizontal motions. In the upper, isothermal layer vertical motions can propagate upward in the form of a pure acoustic wave along the magnetic field lines. Since the sound speed is constant in the upper layer, there is no downward reflection of such an acoustic wave, and it will propagate indefinitely upward with growing amplitude due to decreasing density. This provides a mechanism for escape of energy, which means we are not dealing with a perfect resonator. For horizontal forcing at our chosen frequency and wavenumber, however, only a small fraction of the energy is converted into acoustic energy that escapes to $z = \infty$. Figure 5 shows the distribution of total energy density (Equation (14)) with height for the lowest resonance. The normalized curve is identical for horizontal and vertical forcing. For positive z the energy density drops sharply at first and then levels off rather abruptly at $z \sim 5H_o$ to a nearly constant value of about 2% of the peak value. By considering individual terms in the expression for total energy, Equation (14), we have established that nearly all the energy near $z = 0$ is in the

form of magnetic energy and kinetic energy due to horizontal motion, whereas for $z \gtrsim 5H_o$ the energy is nearly all due to vertical motions, buoyancy, and compression. The sharp drop in energy density between $z = 0$ and $z = 5H_o$ is due to strong downward reflection by the exponentially growing Alfvén speed. The small, constant energy density above $z = 10H_o$ represents escaping energy in the form of a pure acoustic wave along the vertical magnetic field lines.

For forcing at the resonant frequency, the velocities and energy density in layer 1 ($z < 0$) are much smaller than in layer 2 ($z > 0$) and are therefore shown separately. Figure 6 shows the horizontal and vertical velocities and the total energy density in the lower layer for both horizontal and vertical forcing at the lowest resonant frequency. For horizontal forcing, the horizontal velocity drops off rapidly with depth and the vertical velocity rises to a maximum at $z \sim -2H_o$ before decaying rapidly as $z \to -\infty$. The purely horizontal, compressive ($k \neq 0$) motion at $z = 0$ induces vertical motions for $z < 0$ through a "squeezing" action. The resulting compressive vertical motions below $z = 0$ have an acoustic character and are reflected strongly upward by the increasing sound speed. For vertical forcing, horizontal motions are produced below $z = 0$ by a similar "squeezing" effect, but to a lesser extent. The total energy in the lower layer is less for the case of vertical forcing.

The buoyant force is not important in causing upward

reflection in layer 1. Solutions for the case $\alpha = (\gamma-1)/\gamma = 0.4$, corresponding to an adiabatic temperature gradient (neutral stability) in layer 1, have the same character as the solution in Figure 6, for which $\alpha = 0.5$ (slightly superadiabatic).

For forcing at the lowest resonant frequency, the total energy density in the lower layer is six orders of magnitude smaller than in layer 2, and hence doesn't even appear in Figure 5. The resonant modes are almost totally confined to layer 2. To illustrate the sharpness of the resonant response, it is useful to look at the response to forcing at a frequency slightly off resonance. Figure 7 shows the distribution of total energy density for both horizontal and vertical forcing at frequency $\Omega = .743$, about 4% above the lowest resonant frequency. For horizontal forcing, the normalized energy density distribution in the upper layer is nearly identical to that of the resonant response (Figure 5), but the actual numerical values are six orders of magnitude smaller and there is comparable total energy in layer 1 ($z < 0$). For vertical forcing at this off-resonance frequency, the energy density only drops to 36% of its peak value as $z \to \infty$, and there is considerably more outward leakage of energy than in the resonant response. In this case, we expect the presence of the corona (layer 3) to have an important influence. For $z \gtrsim 5H_o$, the energy density is almost completely due to acoustic-like motions along the magnetic field lines; such motions will be reflected strongly at the chromosphere-corona transition.

4.2 THREE-LAYER MODEL

In Section 4.1 above we have established that there is a resonant response of the umbral atmosphere, at a frequency appropriate to umbral oscillations, in a two-layer model atmosphere without a corona. The necessary downward reflection is provided by the rapidly increasing Alfvén speed in layer 2 (the photosphere and chromosphere). Indeed, the energy density of the resonant response has dropped to 2% of its peak value at the assumed height of the transition region, $z_t = 20\ H_o$ (Figure 5). Thus, for umbral oscillations, we find the presence of the high-temperature corona to be of only minor importance. This is in sharp contrast with Uchida and Sakurai (1975) who place heavy emphasis on downward reflection of energy at the transition region.*

For completeness, we consider the effect of the chromosphere-corona transition region on the lowest resonant response. We have calculated the response to horizontal and vertical forcing in the full three-layer model atmosphere (Section 2), with the transition region at $z_t = 20\ H_0$. The solution in layer 3 (the corona) is subjected to the outward radiation condition (15), with H_o replaced by H_1 at $z = 40\ H_o$. Figure 8 shows the distribution of total energy density, which is the same for both horizontal and vertical

*Their emphasis on reflection at the transition region is unwarranted, however; an investigation of their analytical solution shows that almost all of the reflection occurs below the transition region, consistent with our model.

forcing. A comparison of Figures 5 and 8 shows that the small fraction of wave energy at $z = z_t = 20\ H_o$ is almost totally reflected downward by the sharp temperature rise. The energy density for $z > 20\ H_o$ is three orders of magnitude smaller than at $z \lesssim 20\ H_o$ and doesn't show up on the scale of these graphs. The downward reflection at $z = z_t$ also sets up a standing wave pattern that leads to the small oscillations in energy density with height above $z = 5\ H_o$, visible in Figure 8.

Downward reflection at the transition region is of more importance in the case of forcing at off-resonance frequencies, where a higher fraction of the peak wave energy penetrates to the height of the transition.

5. Free Eigenmodes

The results of our investigation of the resonant response of the umbral atmosphere to forcing at $z = 0$ allow us to take a simpler approach to the problem. We have seen that the forced resonant modes are strongly reflected in layer 1 ($z < 0$), with negligible wave energy in that layer. Thus, at least as far as the resonant modes are concerned, it is reasonable to replace layer 1 by a rigid, perfectly conducting lower boundary at $z = 0$. Uchida and Sakurai (1975) assume such a lower boundary in their analysis of quasi-Alfvén waves; however, their stated justification for this is quite different than ours. Uchida and Sakurai argue that the region below $z = 0$

acts as an almost rigid boundary because of the higher density there. However, Thomas (1978) has shown that pure Alfvén waves are only weakly reflected when propagating downward into a region identical to layer 1; thus, Uchida and Sakurai's reasons for assuming a rigid lower boundary are questionable. In our present calculations with the complete wave equations, including the effects of compression, we find that it is the compressive nature of the wave motions and the increasing temperature (and sound speed) with depth that lead to strong upward reflection from layer 1.

Our results for forced resonances also show that there is strong downward reflection in the isothermal layer (layer 2), with only a very small fraction of the wave energy reaching as high as the chromosphere-corona transition region. This suggests that, instead of an outward radiation condition at some height, we could simply assume another perfectly reflecting boundary at some sufficiently large height $z = h$ without changing the nature of the wave mode appreciably.

As a simplified model of umbral oscillations, then, we can look for free eigenmodes of oscillation in an isothermal layer (layer 2) bounded above and below by perfectly reflecting rigid boundaries. We seek solutions to the wave equations (2), (3) in the isothermal region $0 \le z \le h$, subject to the boundary conditions

$$\hat{u} = \hat{w} = 0 \quad \text{at} \quad z = 0, h \ .$$

We expect that the calculated eigenmodes will be fairly independent of h for sufficiently large h ($h \gtrsim 10\ H_o$ for the lowest mode) because most of the downward reflection will have occurred below $z = h$.

Our method of solution is to choose a value of h and calculate the eigenmodes and eigenfrequencies Ω_E using our numerical procedure. The calculations are then repeated for a larger value of h to assure that we have achieved the desired accuracy (Ω_E to three significant figures). The resulting eigenfrequencies Ω_E for the first three modes are listed in Table 1 for comparison with the resonant frequencies Ω_R determined from the forced solutions. As expected, the agreement between Ω_R and Ω_E is good. The computed eigenmodes resemble very closely the resonant modes shown in Figure 3.

An even simpler analytical approach to the free eigenmodes is possible. The nature of the eigenmodes is most closely associated with the horizontal motions, with the vertical motions playing a passive role. This suggests that we can neglect the terms involving the vertical velocity $\hat{w}$ in the equation for the horizontal velocity $\hat{u}$, Equation (2). This gives a closed equation for $\hat{u}$, and the vertical motion $\hat{w}$ is obtained by solving Equation (3) with the known $\hat{u}(z)$. This is the "quasi-Alfvén" approximation of Uchida and Sakurai (1975), which is also used by Hollweg (1979). These authors justify the approximation solely on the basis that

$v_A^2 >> c^2$, which is not true in the lower part of the atmosphere where much of the energy of the eigenmodes resides; in fact, $v_A^2 < c^2$ at $z = 0$ in both Uchida and Sakurai's and our umbral model. The real justification for the approximation low in the atmosphere is that the actual computed values of the terms involving $\hat{w}$ in Equation (2) are much smaller than terms involving $\hat{u}$ for the computed eigenmodes. High in the atmosphere the approximation is valid on the basis that $v_A^2 >> c^2$.

Under the quasi-Alfvén approximation, Equation (2) reduces to

$$\frac{v_A^2}{c_0^2}\frac{d^2\hat{u}}{dz^2} - \frac{K^2}{H_o^2}\left(\frac{v_A^2}{c_o^2} - 1\right)\hat{u} + \frac{\Omega^2}{H_o^2}\hat{u} = 0 \tag{18}$$

in an isothermal layer where $v_A^2 = v_{Ao}^2 \exp(z/H_o)$. The solution to this equation satisfying the conditions $\hat{u}(0) = 0$ and $\hat{u} \to 0$ as $z \to \infty$ is given in terms of Bessel functions in the form

$$\hat{u}_n(z) = J_{2K}\left[\frac{2c_o}{V_{Ao}}(\Omega_n^2 - K^2)^{\frac{1}{2}}\exp(-z/2H_o)\right], \tag{19}$$

where

$$\Omega_n = \left(\frac{1}{4}\frac{v_{Ao}^2}{c_o^2} j_{2K,n}^2 + K^2\right)^{\frac{1}{2}} \tag{20}$$

and where $j_{2K,n}$ is the nth zero of $J_{2K}(z)$. We consider the eigenfrequencies Ω_n as approximate values of the

eigenfrequencies Ω_E computed using the full wave equations and denote them by the symbol $\tilde{\Omega}_E$. Values of $\tilde{\Omega}_E$ for the lowest three eigenmodes are given in Table 1. The values of $\tilde{\Omega}_E$ are indeed close to the corresponding values of Ω_E and, coincidentally, even closer to the corresponding values of Ω_R. Thus, the simple quasi-Alfvén solution (19), (20), which is similar to that of Uchida and Sakurai (the upper boundary condition is slightly different), gives a good estimate of the oscillation period even though certain details of the wave motion are lost.

6. Conclusion and Discussion

We identify umbral oscillations with the lowest resonant mode of fast magneto-atmospheric wave in our model umbral atmosphere. The trapping of this resonant mode is due primarily to the increasing Alfvén speed upward into the chromosphere and the increase in sound speed downward into the convection zone; thus, both the magnetic and acoustic nature of the wave are quite important, with buoyancy playing a smaller role. This trapping mechanism is similar to that proposed for running penumbral waves by Nye and Thomas (1974, 1976), for the case of a horizontal magnetic field. In the present case, the downward reflection is not complete; a small fraction of the total energy in the mode escapes to large heights by converting into the form of an acoustic wave along the vertical magnetic field lines. This small

energy leak may be important in heating the corona above an active region. Downward reflection from the corona is unimportant in determining the character of umbral oscillations, since most of their energy is reflected well below the transition region.

The period of umbral oscillation in our model, for our assumed values of the parameters (Section 2), is 153 s. If we change only the value of the magnetic field strength, we would of course change this period. However, we accept the convincing argument of Uchida and Sakurai (1975) that overstable convection occurs only at a level where the Alfvén speed is not too different from the sound speed. Above or below this level the magnetic field lines are too stiff or too weak, respectively, for overstable oscillation to occur. This means that, regardless of the field strength of the umbra, the value of the Alfvén speed v_{Ao} at the level of forcing z = 0 is roughly the same. According to Uchida and Sakurai, this accounts for the fairly narrow range of observed periods of umbral oscillation among spots of widely differing field strength. In terms of our model, it means that we represent different sunspots by changing the field strength B_o and the level of forcing z = 0 simultaneously, with v_{Ao}, and hence the period of oscillation, nearly unchanged. The observed spread of oscillation periods is more likely due to variations in the size of sunspots, that is, variations in the horizontal wavenumber k.

Finally, a comment should be made concerning the possible cooling of sunspots by magnetohydrodynamic waves. One of us (Thomas 1978) has shown that pure Alfvén waves ($k = 0$) are reflected strongly downward in the umbral photosphere, but are only weakly reflected upwards in the umbral convection zone. This left open the possibility of cooling by downward - propagating waves. However, it was speculated that when more realistic magneto-atmospheric waves, including the effects of compression, are considered, the upward reflection would be much stronger. The results for layer 1 in the present paper confirm this. The upward reflection of waves in layer 1 is very strong for the horizontal wavenumber k considered here and also for higher values of k. This, together with the results of Thomas (1978), argues strongly against significant cooling of sunspots by waves.

ACKNOWLEDGEMENTS

We are grateful to Hugh Van Horn and Don Winget for the use of their stellar pulsation code, which served as the basis for our computer program. We had several useful discussions of this work with Alfred Clark, Jr. This research was sponsored by NASA under grant NSG-7562 and by the Air Force Geophysics Laboratory under contract F19628-77-C-0079.

REFERENCES

Antia, H.M. and Chitre, S.M.: 1978, Solar Phys. 63, 67.

Beckers, J.M. and Schultz, R.B.: 1972, Solar Phys. 27, 61.

Bhatnagar, A. and Tanaka, K.: 1972, Solar Phys. 24, 87.

Ferraro, V.C.A. and Plumpton, C.: 1958, Astrophys. J. 127, 459.

Giovanelli, R.G.: 1972, Solar Phys. 27, 71.

Hollweg, J.V.: 1979, Solar Phys. 62, 227.

Moore, R.L.: 1973, Solar Phys. 30, 403.

Moore, R.L. and Tang, F.: 1975, Solar Phys. 41, 81.

Mullan, D.J. and Yun, H.S.: 1973, Solar Phys. 30, 83.

Nye, A.H. and Thomas, J.H.: 1974, Solar Phys. 38, 399.

Nye, A.H. and Thomas, J.H.: 1976, Astrophys. J. 204, 582.

Phillis, G.L.: 1975, Solar Phys. 41, 71.

Rice, J.B. and Gaizauskas, V.: 1973, Solar Phys. 32, 421.

Thomas, J.H.: 1978, Astrophys. J. 225, 275.

Uchida, Y. and Sakurai, T.: 1975, Publ. Astron. Soc. Japan 27, 259.

TABLE 1

The nondimensional resonant frequency Ω_R for forced oscillations (Section 4.1), the exact free eigenfrequency Ω_E (Section 5), and the approximate free eigenfrequency $\tilde{\Omega}_E$ (Section 5), for the lowest three modes (n=1,2,3) in each case. The dimensional periods τ_R, corresponding to the values of Ω_R, are based on the values $T_o = 4500$ K, $B_o = 1000$ G.

n	Ω_R	Ω_E	$\tilde{\Omega}_E$	τ_R (s)
1	0.712	0.743	0.713	153
2	1.51	1.52	1.50	72
3	2.33	2.46	2.29	47

Fig. 1. The three-layer model of the umbral atmosphere.

Fig. 2. Peaks energy density E_{max} as a function of the forcing frequency Ω for horizontal forcing in the two-layer model atmosphere. Each plotted point represents the result of a numerical solution. The numerical values of E_{max} are based on the forcing amplitude $\hat{u}(0)=1$ km s^{-1}.

Fig. 3. The normalized horizontal velocity as a function of height in layer 2 for the lowest three resonant frequencies in the two-layer model atmosphere. Each higher mode has an additional node and all the profiles have the same asymptotic behavior for large z.

Fig. 4. The logarithmic derivative of the vertical velocity as a function of height in layer 2 for the lowest resonant frequency in the two-layer model atmosphere.

Fig. 5. The total energy density E as a function of height in layer 2 for either horizontal or vertical forcing at the lowest resonant frequency in the two-layer model atmosphere.

Fig. 6. Horizontal and vertical velocity profiles and the total energy density distribution in layer 1 for horizontal forcing (a,c) and vertical forcing (b,d) at z=0. The horizontal velocity $\hat{u}$ is the solid curve and the vertical velocity $\hat{w}$ is the dashed curve.

Fig. 7. The total energy density E as a function of height in layer 2 for horizontal (solid curve) and vertical (dashed curve) forcing at Ω=0.743, about 4% above the lowest resonant frequency, in the two-layer model

atmosphere.

Fig. 8. The total energy density E as a function of height in layers 2 and 3 for either horizontal or vertical forcing in the three-layer model atmosphere. The energy density drops three orders of magnitude across the transition region $z=z_t=20\ H_o$ and is not visible above $z=z_t$ at this scale.

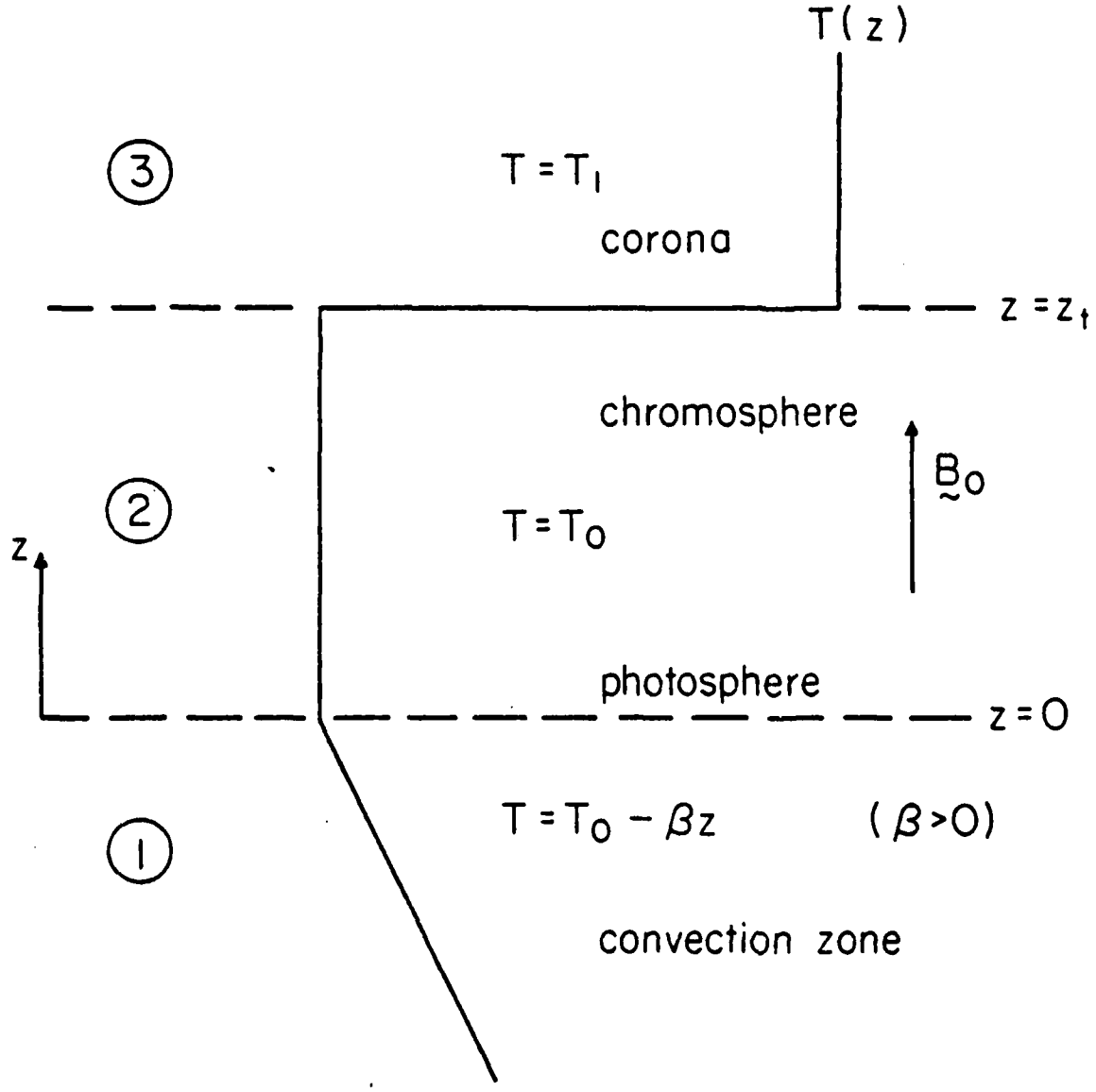

T(z)
3
T = T1
corona
z = zt
chromosphere
B0
2
T = T0
z
photosphere
z = 0
1
T = T0 − βz
(β>0)
convection zone

Figure 1

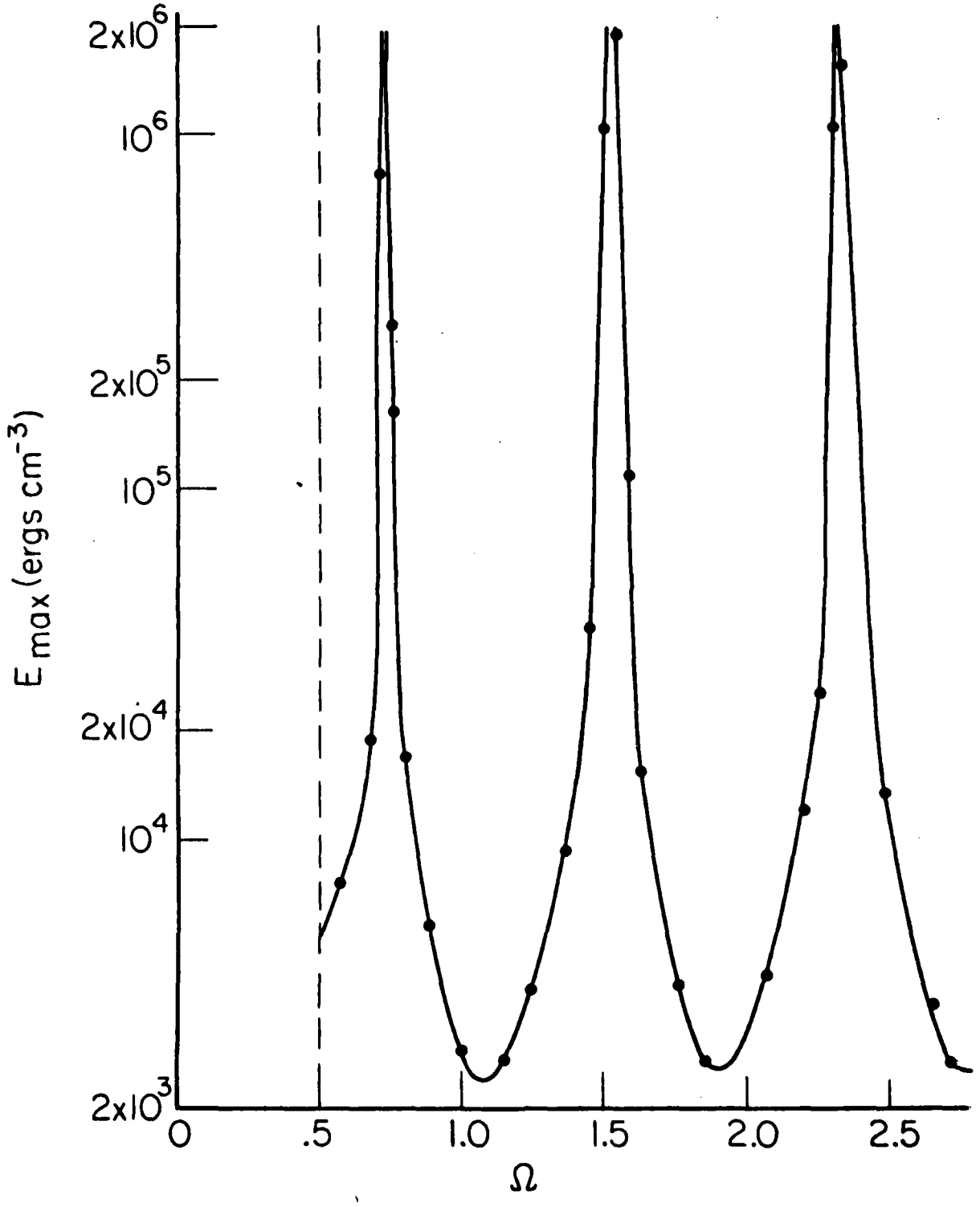

2x10^6
10^6
2x10^5
10^5
2x10^4
10^4
2x10^3
E max (ergs cm^-3)
0
.5
1.0
1.5
2.0
2.5
Ω

Figure 2

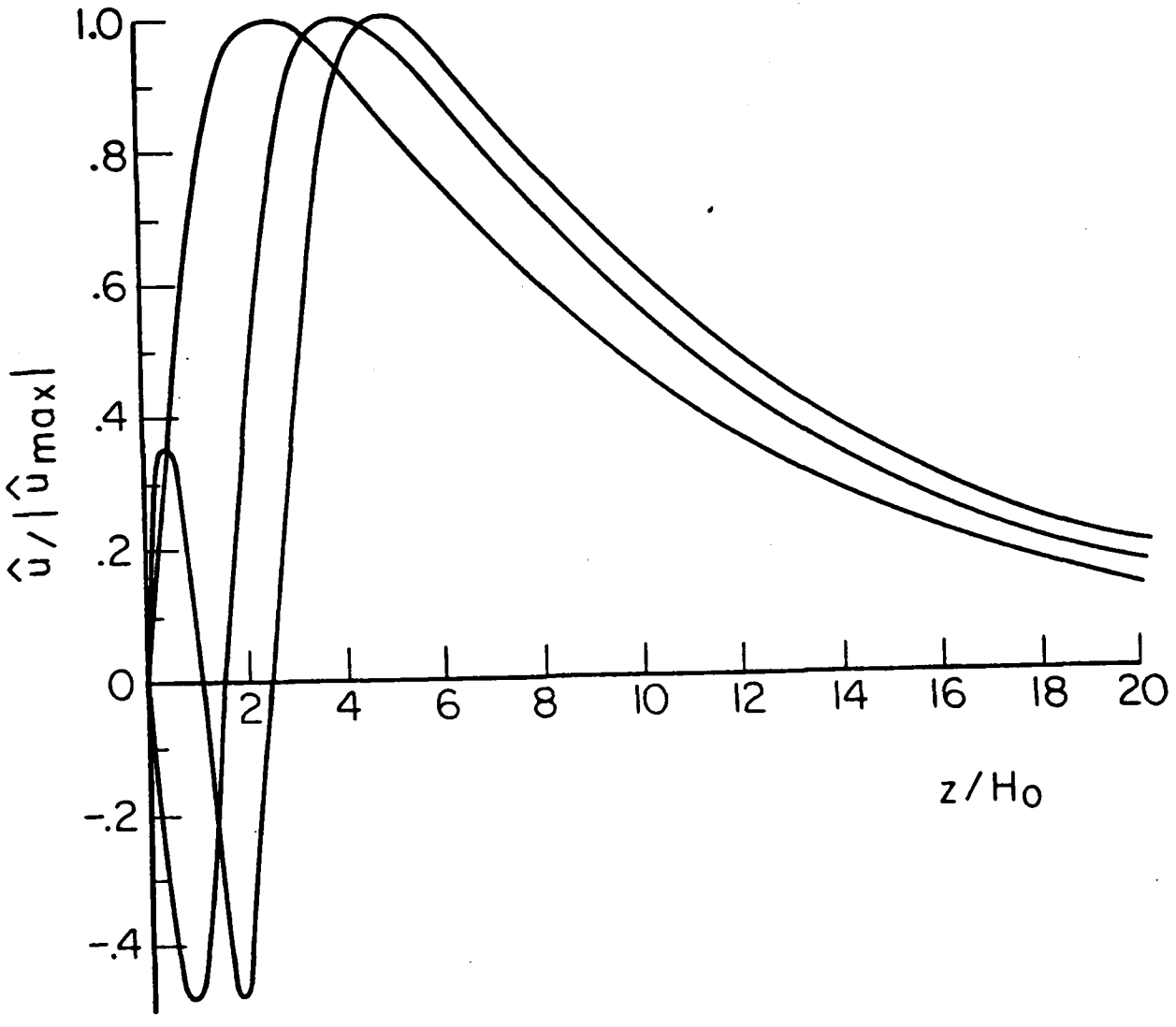


Figure 3

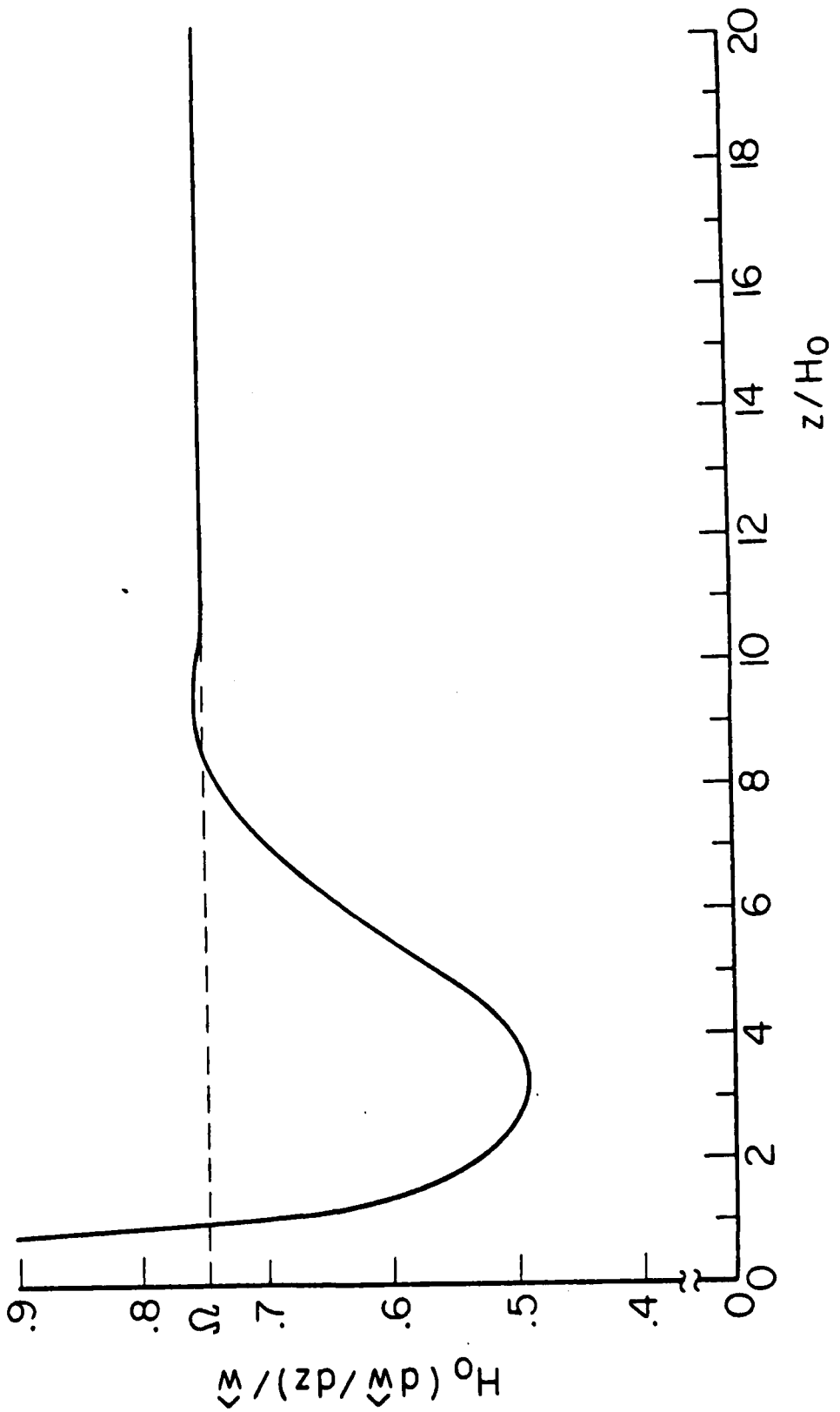


Figure 4

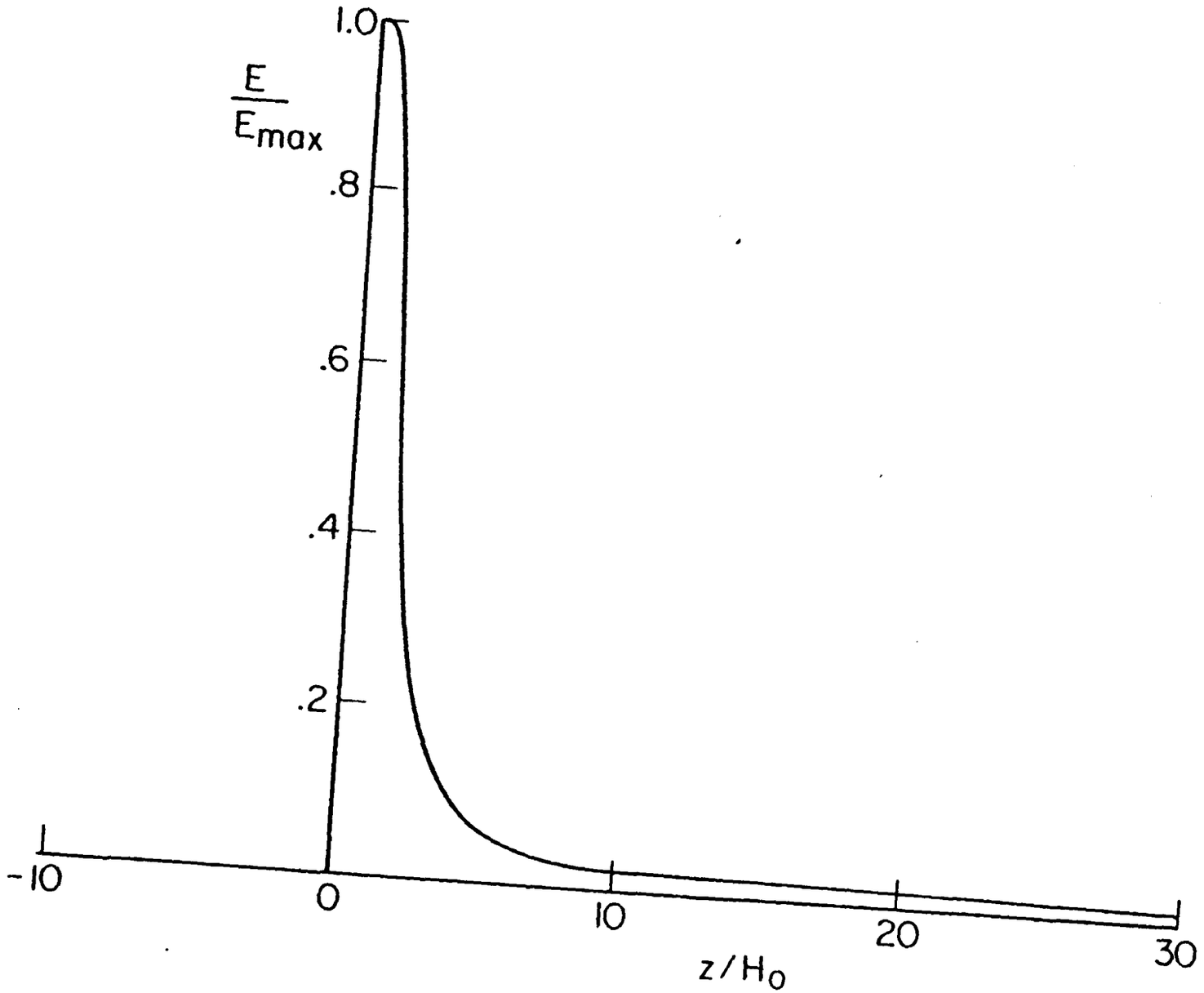


Figure 5

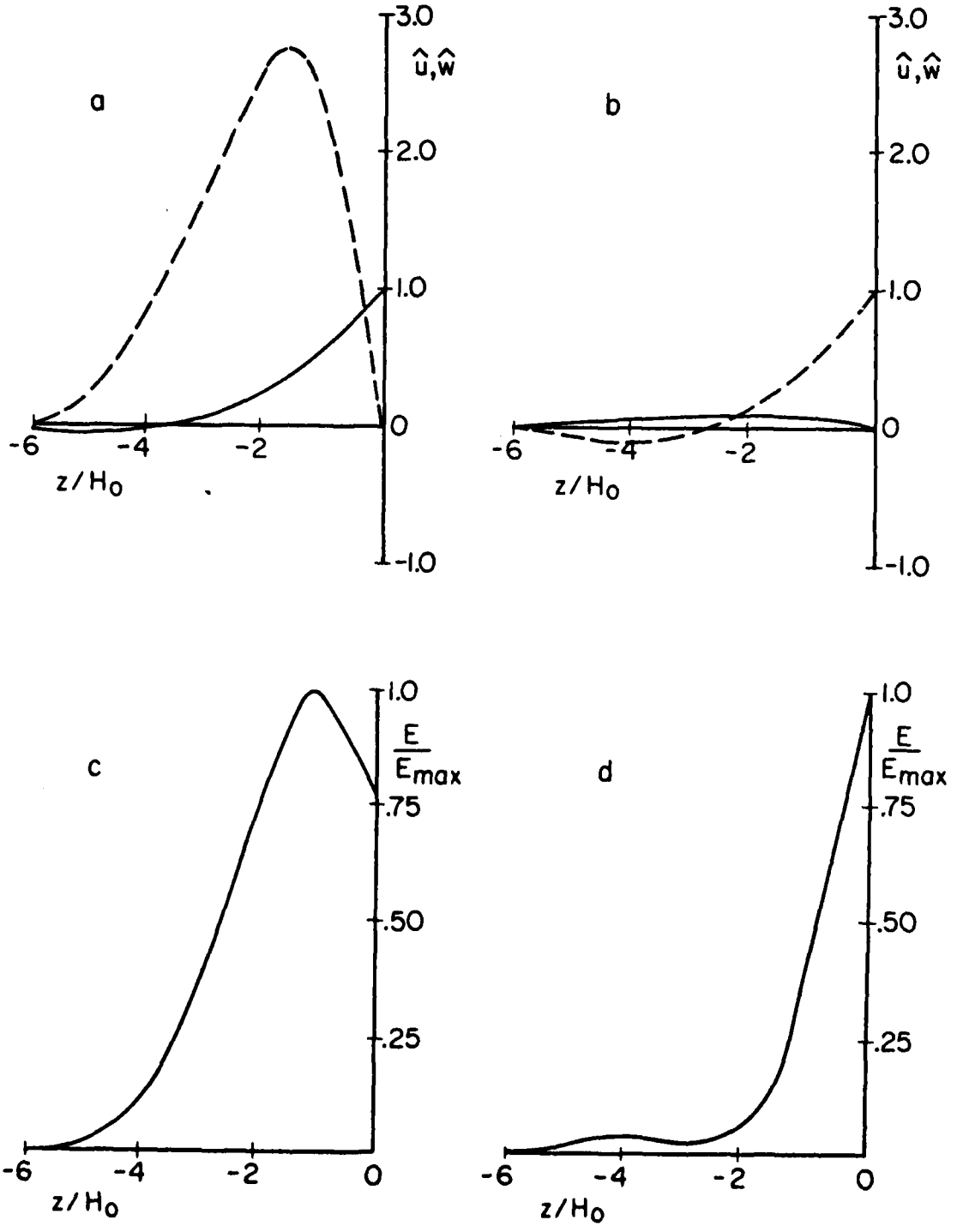


Figure 6

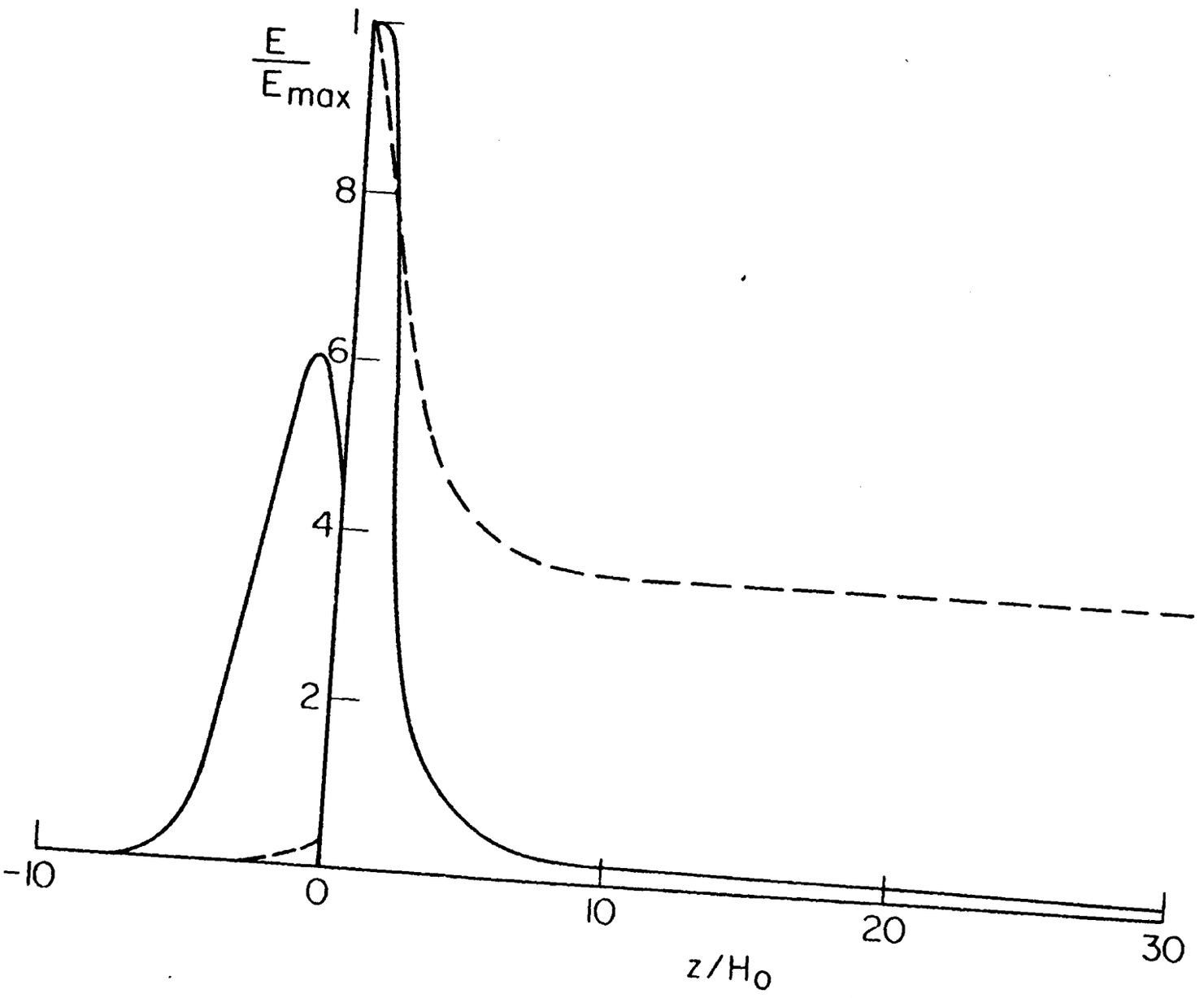


Figure 7

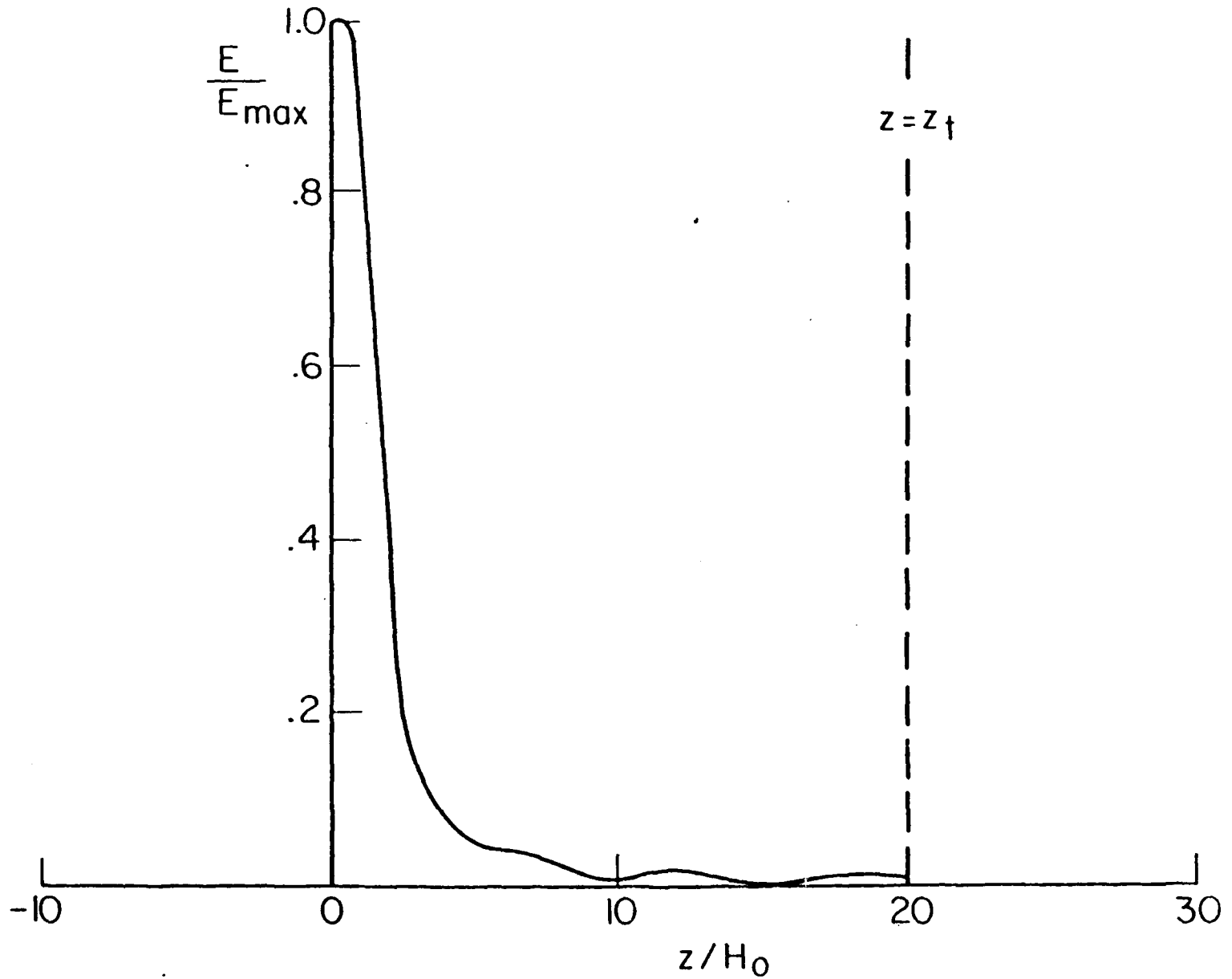


Figure 8


DAT
FILM